A comparative analysis of the intrauterine transcriptome in fertile and subfertile mares using cytobrush sampling

Katharina S. Weber 1, Karen Wagener 2, Miguel Blanco 3, Stefan Bauersachs 4,*,# and Heinrich Bollwein 1,*,#

1 Clinic of Reproductive Medicine, Department for Farm Animals, University of Zurich, AgroVet-Strickhof, 8315 Lindau (ZH), Switzerland; kweber@vetclinics.uzh.ch (K.S.W.); hbollwein@vetclinics.uzh.ch (H.B.)

2 Clinical Unit for Herd Health Management in Ruminants, Department for Farm Animals and Veterinary Public Health, University of Veterinary Medicine, Vienna, Austria; Karen.Wagener@vetmeduni.ac.at

3 The Lewitz Stud, Neustadt-Glewe, Germany; miguelblanco_vet@hotmail.com

4 Institute of Veterinary Anatomy, Vetsuisse Faculty Zurich, University of Zurich, AgroVet-Strickhof, 8315 Lindau (ZH), Switzerland; stefan.bauersachs@uzh.ch

* Correspondence: hbollwein@vetclinics.uzh.ch (H.B.); stefan.bauersachs@uzh.ch (S.B.); Tel.: +41-52-354 91 28 (H.B.); +41-52-354 91 85 (S.B.)

# equal contribution

Abstract:

Background: Subfertility is a major problem in modern horse breeding. Especially, mares without clinical signs of reproductive diseases, without known uterine pathogens and no evidence of inflammation but not becoming pregnant after several breeding attempts are challenging for veterinarians. To obtain new insights into the cause of these fertility problems and aiming at improving diagnosis of subfertile mares, a comparative analysis of the intrauterine transcriptome in subfertile and fertile mares was performed. Uterine cytobrush samples were collected during estrus from 57 mares without clinical signs of uterine diseases. RNA was extracted from the cytobrush samples and samples from 11 selected subfertile and 11 fertile mares were used for Illumina RNA-sequencing.

Results: The cytobrush sampling was a suitable technique to isolate enough RNA of high quality for transcriptome analysis. Comparing subfertile and fertile mares, 114 differentially expressed genes (FDR = 10%) were identified. Metascape enrichment analysis revealed that genes with lower mRNA levels in subfertile mares were related to ‘extracellular matrix (ECM)’, ‘ECM-receptor interaction’, ‘focal adhesion’, ‘immune response’ and ‘cytosolic calcium ion concentration’, while DEGs with higher levels in subfertile mares were enriched for ‘monocarboxyl acid transmembrane transport activity’ and ‘protein targeting’.

Conclusion: Our study revealed significant differences in the uterine transcriptome between fertile and subfertile mares and provides leads for potential uterine molecular biomarkers of subfertility in the mare.

Keywords: Mare, subfertility, uterine transcriptome, cytobrush, RNA-seq, biomarker

1. Background

Subfertility represents a substantial problem for the horse breeding industry (1) as it leads to high economic losses for the owners. Subfertile mares do either not conceive or require more examinations, inseminations and treatments to get pregnant than their fertile counterparts. Many factors such as age, reproductive status, gynecological health of the mare, sperm quality, sperm preservation and breeding management have an effect on fertility (2-4). Clinical endometritis is one of the most common causes for fertility problems in mares (1) and was ranked in the top three medical problems in equine adult patients (5). Endometritis can be divided into acute infectious, chronic infectious or non-infectious endometritis. The most common types of endometritis in mares are bacterially infectious endometritis and persistent breeding induced endometritis (PBIE) (6, 7). Mares susceptible to PBIE show prolonged persistent post breeding uterine inflammation, interfering with the arrival of the embryo in the uterus 5–6 days after breeding (8). Mares with endometritis have a lower conception rate and a higher risk for early embryonic death and mid-gestational abortion. Clinical signs of endometritis include intrauterine fluid, excessive pattern of endometrial edema, vaginitis, vaginal discharge, abnormal estrous cycles and cervicitis. Often endometritis can be diagnosed by detecting clinical signs, uterine inflammation in cytological examination or pathogens in uterine microbial culture (9). However, there are also mares which don't get pregnant after several breeding attempts with sperm of fertile stallions without showing any pathological signs using these diagnostic methods. Le Blanc and Causey (9) described these disturbances in fertility as hidden cases of endometritis or subclinical endometritis.

Although in many studies the histological examination of uterine biopsy samples was considered as the gold standard for diagnosis endometritis (10-13) and for predicting fertility by using the Kenney and Doig score (14), in practice, currently mostly double-guarded uterine swabs for microbial culture and cytobrushes for cytology are used, as these methods are less invasive than the biopsy and less time consuming than histological examination. The sensitivity of microbial culture and cytology is low and these diagnostic methods have a high incidence of false negative results (6, 10, 13, 15). Many bacteria are difficult to cultivate *in vitro* and are therefore not detectable by classical bacteriology (16-18). Moreover, some bacteria, e.g. gram negative bacteria like *Escherichia coli* don't induce a cellular immunological reaction with a high amount of neutrophils detected by the cytological examination in contrast to other bacteria, such as *Streptococci* (1, 19). Therefore, for mares without clinical signs of uterine diseases, without known pathogens in culture, no evidence of inflammation in cytology but not becoming pregnant after several breeding attempts more accurate diagnostic methods are needed to predict fertility.

It seems likely that underlying mechanisms for subfertility can be found at the molecular level. For instance mares susceptible to persistent endometritis show differences in innate immune response to insemination (8, 20-22) and induced infectious endometritis (23) compared to resistant mares at mRNA expression level. The mRNA expression of pro-inflammatory cytokines (*IL6, IL1RN, IL1B, IL8*), anti-inflammatory cytokines (*IL10*), tumor necrosis factor (*TNFα*), C-C motif chemokine ligand 2 (*CCL2*), antimicrobial peptides secreted phospholipase A2 (*sPLA2*), lipocalin 2 (*LCN2*) and lactoferrin (*LFN*) differ between susceptible and resistant mares (8, 20-23). Recently, it has been shown that mares susceptible for PBIE show a different expression pattern of genes associated with innate immunity even before breeding and that antimicrobial peptides equine b-defensin 1 (*EBD1*), lysozyme (*LYZ*) and secretory leukoprotease inhibitor (*SLPI*) can be used as diagnostic marker for susceptibility (22).

Gene expression profiling of the healthy, receptive equine endometrium has shown that the transcriptome differed among estrous cycle stages (24, 25). Genes upregulated during

estrus were associated with extracellular matrix related categories and immune regulated functions (24, 25). These physiological changes in uterine gene expression could play an important role in successful reproduction. For instance, the uterine immune system may prepare the uterus for potential foreign material ascending through the open cervix during estrus by upregulation of genes related to immune response (24, 25).

In recent years, gene expression analysis has been applied in several studies to identify genes and their networks associated with receptivity of the human endometrium at the time of implantation by comparing women with recurrent miscarriage (26, 27) or recurrent implantation failure (26, 28-30) and fertile women. Furthermore, in cows, different studies were performed to identify endometrial gene expressions related to fertility (31-35). However, to our knowledge, no study investigated the relationship between the equine uterine transcriptome and fertility yet.

In most of the equine and human studies uterine biopsy samples were taken for transcriptome and mRNA analysis, while in cattle cytobrush samples were often used for mRNA analysis. In different bovine studies, it was shown that cytobrush sampling provides a much less invasive method to isolate RNA of sufficient quantity and quality for gene expression analysis (31, 36) compared to the biopsy of the endometrium.

With the aim to improve the diagnosis of subfertile mares without clinical signs of uterine diseases and to characterize RNA markers to predict fertility, our objective was to perform a comparative analysis of the intrauterine transcriptome at estrus of fertile and subfertile mares without clinical signs of uterine diseases. A second objective was to investigate the suitability of samples collected by cytobrush from the equine uterus for transcriptome analysis.

2. Methods

2.1 Mares

The study was conducted between July and October 2018 at a large private commercial stud farm in Mecklenburg-Vorpommern, Germany. In total, 57 standard bred mares aged between 5 and 14 years presented for routinely bacteriological examination before artificial insemination (AI) were selected for the study. Mares with foal as well as barren and maiden mares were represented. To avoid subfertility because of age related modifications of the endometrium only mares until age of 14 were included. All mares showed no clinical signs of reproductive disorders and of PBIE after previous inseminations or had no history of being susceptible to PBIE. From these 57 mares 11 fertile and 11 subfertile mares were selected for comparative transcriptome analysis.

2.2 Study design and selection of fertile and subfertile mares

The 57 mares were initially subdivided according to their breeding history of the breeding season 2018 in mares having their first insemination this year (First Breeders (FB), n = 30) and in mares that failed to conceive after at least two consecutive AI (Repeat Breeders (RB), n = 27; Fig. 1). During estrus, the reproductive tract of the mares was examined daily by transrectal palpation and ultrasonography (Aloka prosound 2, Hitachi, Japan) to assess the optimal time for sampling and AI and to examine the mares for clinical signs of endometritis including intrauterine fluid, an excessive pattern of endometrial edema, vaginitis, vaginal discharge, abnormal estrous cycles and cervicitis (9, 37). Only mares with no clinical signs for endometritis or reproductive abnormalities were included in the study. Uterine samples were collected as described later from mares in midestrus (dominant follicle >35 mm, distinct endometrial edema). Ovulation was induced using human chorionic gonadotropin (hCG, Ovogest 1500 I.U. i.v.) when the follicle reached a diameter of >40 mm and AI was

performed 24 h later. To avoid subfertility because of low sperm quality, only chilled semen from known fertile stallions having a sperm motility >50% were taken. Twenty-four hours after AI, the mares were examined for ovulation and intrauterine fluid by transrectal ultrasound. The first pregnancy diagnosis was carried out 16 days after single-ovulations or 14 days after double-ovulations by transrectal ultrasound. A second pregnancy diagnosis was performed 45 (+/-2 d) days after ovulation. Mares that were positive in both pregnancy examinations were counted as pregnant. According to the results of the pregnancy examinations, mares were finally grouped into FB and RB pregnant (FB-P and RB-P) and not pregnant (FB-N and RB-N). FB-P were considered fertile (n = 19) and RB-N were defined subfertile (n = 12; Fig. 1). Mares with a positive cytological finding or facultative pathogenic bacteria in microbial culture were excluded from further analysis. Eleven mares of the subfertile group (RB-N) with no facultative pathogenic bacteria and negative cytology were used for transcriptome analysis. According to these mares, 11 fertile mares with similar age and reproductive status, negative cytology and without facultative pathogens in microbial culture were selected from the FB-P group for RNA sequencing.

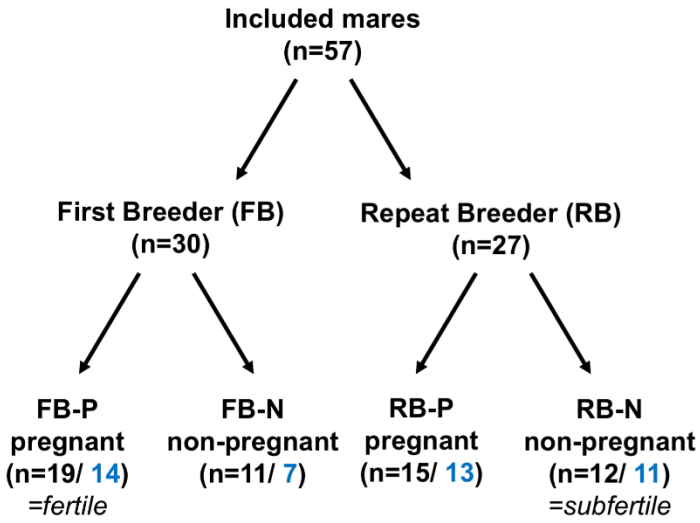

Figure 1. Grouping of mares according to their breeding history and fertility: First Breeder: uterine samples were taken before first AI during the season; Repeat Breeder = mares failed to conceive after at least two consecutive AIs; FB-P = mares pregnant after the first breeding during the season; FB-N = mares not pregnant after the first breeding during the season; RB-N = mares pregnant after the third breeding during the season; RB-N = mares not pregnant after at least three breedings during the season; black numbers indicate the total number of mares in each group; blue numbers indicate the number of mares after exclusion of mares with facultative pathogenic bacteria in uterine swabs. FB-P were defined fertile, RB-N subfertile. Eleven samples of FB-P and 11 samples of RB-N were used for transcriptome analysis.

2.3 Uterine sample collection

From the 57 mares, uterine samples were taken in estrus (dominant follicle >35 mm, distinct endometrial edema). Before sample collection, the rectum was emptied, the tail was wrapped and covered with a glove, the perineum and vulva were cleaned with soap and water and dried with a paper towel. A veterinarian (K.S.W.) wearing a sterile glove inserted manually one double guarded uterine swab (Minitube, Tiefenbach, Germany). After passing the cervix, the inner guard was advanced, and the swab was rolled gently on the endometrium for 15 seconds. The swab was retracted and removed together with the inner guard from the reproductive tract. The outer guard was left in the reproductive tract to avoid additional manipulation during the following sampling steps. The swab was transferred into Amies medium and sent to a laboratory specialized for bacteriology in equine reproduction (Labor Boese, Harsum, Germany) for conventional microbiological culture. A new inner guard with a cytobrush (Minitube, Tiefenbach, Germany) was inserted through the outer guard into the uterus. The brush was advanced and rolled gently on the endometrium for 15 seconds. The brush was retracted, removed from the reproductive tract and was inserted directly into a 1.5 ml reaction tube containing 350 µl of lysis buffer (AllPrep DNA/RNA Kit, Qiagen), rolled for 20 seconds and then removed. The reaction tubes were snap-frozen in liquid nitrogen until

further analysis. In the same way, a second cytobrush sample was collected. The second cytobrush was rolled onto two microscope slides for cytological examination.

2.4 Sample examination

2.4.1 Cytological examination

After air drying, the microscope slides were stained with Diff-Quick (Diff Quick, Labor+Technik Eberhard Lehmann GmbH, Germany). The slides were evaluated under a light microscope (Olympus Ch-2, Olympus, Tokyo, Japan) as described by Ferris (17). Briefly, at 400 x magnification, 10 high power fields were analyzed, and the number of neutrophil granulocytes was counted. The presence of neutrophil granulocytes defines the following categories: no granulocytes to rare is a negative finding (healthy), 1–2 granulocytes per field indicates a mild inflammation, 3–5 moderate inflammation, >5 severe inflammation (17).

2.4.2 Bacteriological examination

Bacteriological examination was performed at laboratory Böse (Harsum, Germany) specialized for bacteriology in equine reproduction. Bacteria were cultured on Columbia agar with 5% sheep blood, Columbia CAP Selective agar with sheep blood and Gassner agar (water blue-metachrome yellow- lactose-agar). Additionally, a mycological examination was performed on Sabouraud-Glucose agar with gentamycin and chloramphenicol and on Kimmig-agar. After incubation at 37° C for 24 hours and 48 hours, the bacterial and fungal growth was evaluated. The aerobic microflora was identified by cultural-biochemical conventional methods and mass spectroscopy (MALDI-TOF). The samples were divided into mild (<30 colonies), moderate (30–100 colonies) and severe (>100 colonies) growth.

The bacteriological results were divided into facultative pathogenic bacteria and questionable pathogenic bacteria. Growth of *ß-hemolytic Streptococcus, Pseudomonas aeruginosa, Klebsiella pneumoniae, Staphylococcus aureus, Candida albicans*, *Aspergillus spp.* and moderate to severe growth of *E. coli* were considered as facultative pathogens. Mild growth of *E. coli* and the growth of other bacteria were considered as questionable pathogens. Samples with facultative pathogens were excluded from further analysis (n = 12).

2.4.3 Extraction of RNA

The total RNA and DNA extraction was performed with the AllPrep DNA/RNA micro kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions using the protocol of Simultaneous Purification of Genomic DNA and Total RNA from Animal and Human Tissues. Because of rolling the cytobrush directly in lysis buffer after taking the sample, the initial start of the protocol was at step three of the protocol. RNA and DNA were extracted according to manufacturer's instructions, instead of step 7 steps F1–F4 were used. RNA was eluted after two five-minute incubations with 14 µl of RNase-free water.

2.4.4 RNA concentration and quality measurement

The total RNA concentration and purity were measured by spectrophotometry (Nano Drop One, Thermo Scientific). The quality of the isolated total RNA was assessed using the Agilent 2100 Bioanalyzer RNA 6000 Nano assay (Agilent Technologies, Waldbronn, Germany).

2.5 RNA-sequencing and data analysis

For Illumina RNA-sequencing, 11 samples of subfertile mares (RB-N) and 11 of the fertile mares (FB-P) were selected. The isolated RNA was used for standard Illumina messenger RNA sequencing following the TrueSeq stranded mRNA protocol, which was performed by the Functional Genomic Center Zürich. The samples were barcoded and pooled and the pool was sequenced on an Illumina NovaSeq 6000 (100 bp single-end reads).

The data analysis of the resulting fastq files was performed on a local Galaxy server installation (38, 39). In Galaxy, as first step and after every fastq file processing step, 'Fast QC' and 'Multi QC' were used to control the quality of the data and the processing steps. At first, 'Trim Galore' was used to trim adapter sequences, low quality ends and to discard reads shorter than 50 bp. To map the reads to the annotated equine genome equCab 3.0 downloaded from the National Center for Biotechnology Information (NCBI), the tool 'HISAT2' (40) was used. With the tool 'Feature Counts', the number of RNA sequence reads for the annotated genes were counted considering the strand specificity of the reads. 'Column Join on data' was used to build a count table containing all samples. This count table was filtered to remove genes with negligible read counts by using the counts per million (CPM) per sample filtering tool (41). The mean library size and potential CPM cutoff (Counttable statistics, custom Galaxy tool) were calculated and the cutoff set to 1.09 CPM (corresponding to an average of 20 reads per library) for at least 5 out of the 22 libraries. The filtered count table was used in the statistical program R with the package edge R (42) for the identification of differentially expressed genes (DEGs) between subfertile and fertile mares. Genes with an adjusted P-value (false discovery rate, FDR) lower than 0.01 (FDR 10 %) were considered as differentially expressed.

The web tool MAdb (Gene Symbol match, Ensembl compara database release 95, Blast) (https://madb.ethz.ch/) (43) was used to obtain the corresponding human gene symbols and gene information for the DEGs. Hierarchical clustering of DEGs was performed with the HCL tool of Multi Experiment Viewer (MeV) (44). Metascape enrichment analysis program (http://metascape.org) (45) was used for identification of overrepresented functional categories and pathways of the DEGs. The annotation and enrichment analyses were performed separately for genes upregulated or downregulated in subfertile mares compared to fertile mares based on the corresponding human NCBI gene IDs.
The RNA-seq data presented in this study are openly available at NCBI's Sequence Read Archive (SRA) under the BioProject accession PRJNA667444.

2.6 Quantitative real-time RT-PCR

The same RNA samples (400 ng total RNA) as used for RNA sequencing were reverse transcribed into first-strand cDNA using the RNA to cDNA EcoDry™ (Double Primed) Premix (Takara Bio Company, USA). The cDNA samples were diluted with RNase/DNase-free water to a total volume of 40 µl.

To validate RNA sequencing results, the relative expression of solute carrier family 10 member 2 (*SLC10A2)*, solute carrier family 16 member 9 (*SLC16A9)*, integrin subunit beta 3 (*ITGB3)*, collagen type 6 alpha 1 chain (*COL6A1)*, thrombospondin 2 (*THBS2)*, fibronectin 1 (*FN1)*, matrix metallopeptidase 25 (*MMP25)*, ectonucleotide pyrophosphatase/phosphodiesterase 3 (*ENPP3)*, atypical chemokine receptor 3 (*ACKR3)* and phosphodiesterase 10 A (*PDE10A)* was determined by quantitative real-time RT-PCR. The mapping of RNA sequencing reads for these genes were checked with Integrated Genomics Viewer (IGV) (46) and matching primers were designed using Primer-Blast (NCBI) (47). The primer sequences (ordered from Integrated DNA Technology, Leuven, Belgium) and their annealing temperatures are listed in Table 1.

The mRNA expression of the selected genes was measured by real-time PCR on a Light Cycler 96 (Roche, Mannheim, Germany) with the KAPA HiFi HotStart PCR Kit (Roche, Kapa Biosystems Pty, South Africa) adding EvaGreen® Dye, 20x in water (Biotium). The qPCR was performed in a reaction volume of 10 µl, consisting of 2 µl 5x Kapa HiFi Buffer, 0.3 µl Kapa dNTP mix (0.3 mM each), 0.2 µl Kapa HiFi HotStart DNA polymerase, 0.3 µl of each primer (10 µM), 0.5 µl Eva Green Dye, 5.4 µl water and 1 µl cDNA template. Cycle parameters of the PCR were 95°C for 3 min, followed by 45 cycles of 98°C for 20 sec, specific annealing temperature for 15 sec and 72°C for 15 sec, and then a melting step (95°C for 10 sec, 65°C for 60 sec and 97°C for 1 sec). Melting curves of the amplified PCR products were obtained for confirmation of specific amplification. A no-template control (RNA sample) was included for each primer pair. The C_q (quantification cycle) values determined for the target genes were normalized against the geometric mean of the reference genes beta actin (*ACTB*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and 18S rRNA (48). Relative expression differences between subfertile and fertile mares were calculated, and a t-test was performed in Microsoft Excel. P-values <0.05 were considered significant.

Table 1. Primers used for quantitative real-time RT-PCR: Ten DEGs were selected to validate RNA sequencing results by qRT-PCR. ACTB, GAPDH and 18S rRNA were used as reference genes. The designed primers are listed with product length and annealing temperature.

Gene	Primer sequence 5′-3′	Product length (bp)	Annealing Temp C°	Accession no./Reference
SLC10A2	F: ATCGTTCACCTACGAGGAGC	191	66	XM_001493450.3
	R: TCACCTTGTGGAGCGATGAC			
SLC16A9	F: TGTTCTTTGCTGGGCTTGGA	110	68	XM_023643589.1
	R: CAGGACGCAGAAGCCACTAA			
ITGB3	F: GCACCCGTTACTGTCGTGAT	145	65	NM_001081802.1
	R: AGGATGGACTTTCCACTGGC			
THBS2	F: TGGCTGGAAAGACTACACCG	107	65	NM_001163117.2
	R: CTGAATCCGCCATGACCTGT			
FN1	F: GGTCGTTACTGTGGGCAACT	101	65	XM_023642280.1
	R: CCTCTCCGATGGCGTAATGG			
COL6A1	F: CCTCCTGGGATAAACGGCAC	184	65	XM_001488351.5
	R: ACTCGTCCATCTCTGGTCGT			
ACKR3	F: ATGCCTGAGTAGCCTGGAGA	113	65	XM_023642191.1
	R: GTCCTGTGGTGATGCAAACG			
LOC100073089	F: TAGAATACGTGGTCAACACCAG	190	68	XM_023651094.1
(ENPP3)	R: TCAACCCAGTTGGCTTCCTG			
PDE10A	F: GCGTGAATTGTAGCAGCCAG	76	65	XM_023633145.1
	R: ACTGATTGCAGAAAGACACTTCC			
MMP25	F: ATGTCACCGTCAGCAACACAG	189	70	XM_023633145.1
	R: GTCCAGGCTTGAGAGTGGCT			
ACTB	F: TCCCAGCACGATGAAGATCAA	189	68	XM_023655002.1
	R: GGTGGATCGCACTAACAGT			
GAPDH	F: ATTGCCCTCAACGACCACTT	140	70	NM_001163856.1
	R: TCTTGCTGGGTGATTGGTGG			
18S rRNA	F: GCGTGTGCCTACCCTACGCC	165	68	AJ311673.1/ (24)
	R: ATCGTTCACCTACGAGGAGC			

3. Results

3.1 Cytology and bacteriology

The cytological examination did not reveal an intrauterine inflammation at the time of sampling in all mares. Bacteria were detected in 33 of 57 mares (57.9%) in microbial culture. Facultative pathogens were obtained in 12 of 57 mares (21.1%). These samples were excluded from further analysis. From each group of the fertile mares (FB-P) and subfertile mares (RB-N) 11 mares without facultative pathogens were selected for RNA sequencing.

3.2 Isolation of RNA from cytobrush samples and Illumina RNA-sequencing

The cytobrush sampling was a suitable technique to isolate enough RNA of high quality for transcriptome analysis. The concentration of the total RNA was between 40 and 669 ng/µl, while the A260/A280 ratio was between 1.95 and 2.09. The obtained RNA integrity numbers (RIN) ranged from 8.9 to 10 in all 57 samples.

The RNA sequencing results revealed after filtering of the fastq files library sizes between 10.9 and 30.9 million reads per sample with an average of 18.4 million reads. After filtering genes with low read counts, in total 15,318 different genes were detectable and used for differential gene expression analysis.

3.3 Identification of differentially expressed genes

The intrauterine transcriptome differed between subfertile and fertile mares without clinical signs of uterine diseases. Using Edge R analysis, 114 genes were found as differentially expressed between subfertile and fertile mares (FDR<0.1; Fig. 2) (Supplementary Data Table S1). Ninety-eight genes were significantly downregulated and 16 genes upregulated in subfertile mares compared to fertile mares. The expression of neuromedin U (*NMU*), synaptogamin 12 (*SYT12*), uncharacterized *LOC111767890*, UL16 binding protein 1 (*LOC100063831*) were decreased to the greatest extent, while the expression of solute carrier family 10 member 2 (*SLC10A2*), 40S ribosomal protein S2- like (*LOC100147232*) and 60S ribosomal protein L26-like (*LOC10052427*) were increased to the greatest extent in subfertile mares compared to fertile mares.

Hierarchical cluster analysis of the DEGs revealed a separation of DEGs upregulated (cluster 1) or downregulated (clusters 2, 3, 4) in samples derived from subfertile mares (Fig. 2). The downregulated genes were separated in three clusters. Cluster 2 showed DEGs with increased expression in only 5 of the fertile mares. Differences in cluster 3 and 4 were more consistent above all samples. A few samples of the subfertile and fertile group, respectively, showed expression patterns in part more similar to the respective other group. The DEGs of cluster 1 and the DEGs of clusters 3 and 4 are listed in Table 2 and Table 3.

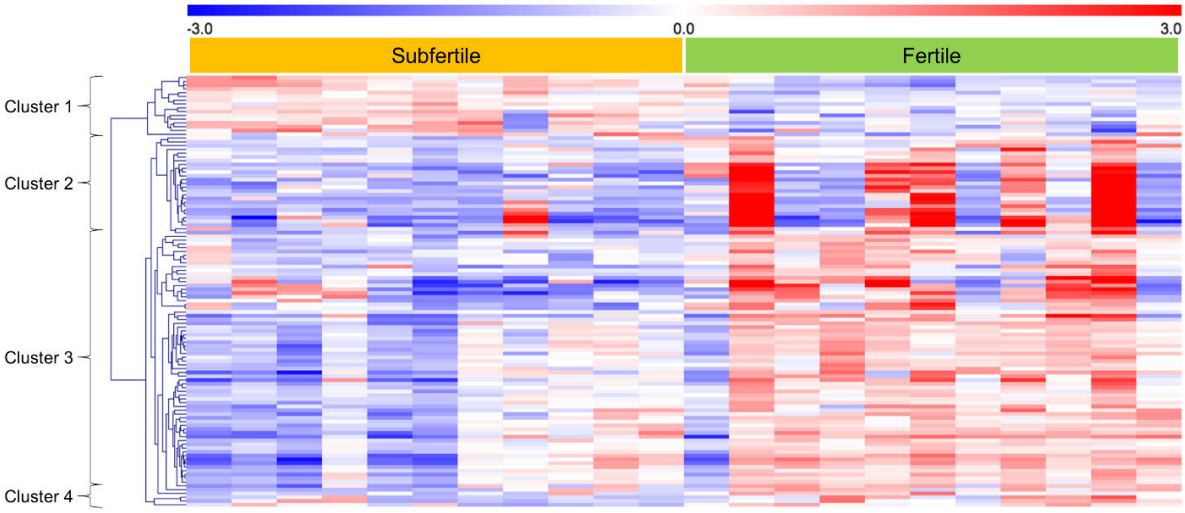

Figure 2. Heat map and hierarchical cluster analysis of DEGs between subfertile and fertile mares (FDR<0.1). Each line represents 1 DEG, each column 1 sample. Red color represents higher and blue color lower expression of the gene compared to the mean of all samples (mean-centered values in log2 scale).

Table 2. DEGs Cluster1: DEGs upregulated in subfertile mares compared to fertile mares.

Gene symbol	Gene ID	Gene description	human Gene symbol	logFC	PValue	FDR
SLC10A2	100051264	solute carrier family 10 member 2	*SLC10A2*	1.82	0.0002	0.0591
LOC100147232	100147232	40S ribosomal protein S2 like		1.80	0.0000	0.0012
LOC100052427	100052427	60S ribosomal protein L26-like		1.64	0.0000	0.0200
LOC100051778	100051778	60S ribosomal protein L21	*RPL21*	1.63	0.0000	0.0035
LOC100065786	100065786	40S ribosomal protein S17	*RPS17*	1.46	0.0000	0.0014
LOC100067178	100067178	Mesothelin	*MSLN*	1.13	0.0005	0.0861
SLC16A9	100062703	solute carrier family 16 member 9	*SLC16A9*	0.98	0.0001	0.0417
ELOVL2	100063624	ELOVL fatty acid elongase 2	*ELOVL2*	0.96	0.0001	0.0474
LOC111767704	111767704	uncharacterized LOC111767704		0.91	0.0000	0.0179
CRYL1	100054141	crystallin lambda 1	*CRYL1*	0.89	0.0001	0.0338
CTSE	100055161	cathepsin E	*CTSE*	0.82	0.0002	0.0591
LOC100072143	100072143	centrin-4		0.80	0.0001	0.0318
SLC16A5	100060017	solute carrier family 16 member 5	*SLC16A5*	0.57	0.0005	0.0861
TOMM7	100630688	translocase of outer mitochondrial membrane 7	*TOMM7*	0.52	0.0002	0.0591
MYCBP	100068904	MYC binding protein	*MYCBP*	0.47	0.0002	0.0564
SDHAF4	100629833	succinate dehydrogenase complex assembly factor 4	*SDHAF4*	0.45	0.0004	0.0763

Table 3. DEGs Cluster 3, 4: DEGs downregulated in subfertile mares compared to fertile mares.

Gene symbol	Gene ID	Gene description	human Gene symbol	logFC	PValue	FDR
ACAP1	100072970	ArfGAP with coiled-coil, ankyrin repeat and PH domains 1	*ACAP1*	-1.01	0.0000	0.0291
ACKR3	100057501	atypical chemokine receptor 3	*ACKR3*	-1.60	0.0001	0.0474
ADAMTS7	100059959	ADAM metallopeptidase with thrombospondin type 1 motif 7	*ADAMTS7*	-1.39	0.0001	0.0318
AKR1E2	100070632	aldo-keto reductase family 1 member E2	*AKR1E2*	-1.93	0.0000	0.0179
ANKRD10	100066458	ankyrin repeat domain 10	*ANKRD10*	-0.66	0.0007	0.0998
ANO8	100146761	anoctamin 8	*ANO8*	-1.24	0.0004	0.0763
APBA3	100146445	amyloid beta precursor protein binding family A member 3	*APBA3*	-0.62	0.0003	0.0708
C1QA	100058097	complement C1q A chain	*C1QA*	-0.66	0.0006	0.0861
C1QB	100071667	complement C1q B chain	*C1QB*	-0.74	0.0002	0.0562
CEP131	100056159	centrosomal protein 131	*CEP131*	-0.85	0.0002	0.0529
CIAO3	100065271	cytosolic iron-sulfur assembly component 3	*CIAO3*	-0.57	0.0007	0.0998
CLK1	100067832	CDC like kinase 1	*CLK1*	-0.94	0.0004	0.0796
CLK2	100063546	CDC like kinase 2	*CLK2*	-0.69	0.0006	0.0861
COL16A1	100056083	collagen type XVI alpha 1 chain	*COL16A1*	-1.72	0.0005	0.0847
COL4A1	100066148	collagen type IV alpha 1 chain	*COL4A1*	-1.16	0.0003	0.0658
COL4A2	100066264	collagen type IV alpha 2 chain	*COL4A2*	-1.17	0.0001	0.0354
COL6A1	100050035	collagen type VI alpha 1 chain	*COL6A1*	-1.84	0.0006	0.0897
CYTH4	100069735	cytohesin 4	*CYTH4*	-0.79	0.0004	0.0798
DENND1C	100065730	DENN domain containing 1C	*DENND1C*	-0.91	0.0006	0.0861
DLG4	100061544	discs large MAGUK scaffold protein 4	*DLG4*	-0.98	0.0003	0.0643
DNASE1L3	100057863	deoxyribonuclease 1 like 3	*DNASE1L3*	-1.66	0.0001	0.0474
EHBP1L1	100057282	EH domain binding protein 1 like 1	*EHBP1L1*	-0.93	0.0001	0.0417
ENTPD6	100057043	ectonucleoside triphosphate diphosphohydrolase 6	*ENTPD6*	-0.70	0.0003	0.0620
FER1L5	100062182	fer-1 like family member 5	*FER1L5*	-1.77	0.0002	0.0591
FN1	100034189	fibronectin 1	*FN1*	-2.21	0.0001	0.0472
GRAMD1B	100063638	GRAM domain containing 1B	*GRAMD1B*	-0.61	0.0004	0.0763
IRF8	100056218	interferon regulatory factor 8	*IRF8*	-0.79	0.0005	0.0861
JAK3	100147451	Janus kinase 3	*JAK3*	-0.66	0.0003	0.0673
KIF7	100069672	kinesin family member 7	*KIF7*	-0.82	0.0006	0.0867
LAT	100064430	linker for activation of T cells	*LAT*	-0.93	0.0001	0.0417
LLGL1	100051856	LLGL1, scribble cell polarity complex component	*LLGL1*	-0.50	0.0006	0.0861
LOC100054029	100054029	leukocyte immunoglobulin-like receptor subfamily A member 5		-1.26	0.0006	0.0883
LOC100054448	100054448	saoe class I histocompatibility antigen, A alpha chain	*HLA-A*	-1.26	0.0000	0.0024
LOC100055483	100055483	Ig mu chain C region membrane-bound form-like	*IGHM*	-1.82	0.0003	0.0609
LOC100063097	100063097	mitotic-spindle organizing protein 2B-like	*MZT2B*	-1.70	0.0001	0.0423
LOC100073089	100073089	ectonucleotide pyrophosphatase/phosphodiesterase family member 3	*ENPP3*	-1.26	0.0000	0.0200
LOC100629324	100629324	uncharacterized LOC100629324	*MEG3*	-2.85	0.0001	0.0417
LOC102149846	102149846	immunoglobulin heavy constant gamma 1-like	*IGHG1*	-2.14	0.0005	0.0861
LOC102150085	102150085	immunoglobulin heavy constant gamma 1-like	*IGHG1*	-2.67	0.0000	0.0240
LOC102150790	102150790	uncharacterized LOC102150790		-1.44	0.0000	0.0179
LOC106781059	106781059	uncharacterized LOC106781059		-1.46	0.0001	0.0327
LOC106781303	106781303	immunoglobulin heavy constant alpha 2-like	*IGHA1*	-1.98	0.0005	0.0861
LOC106781940	106781940	uncharacterized LOC106781940		-1.45	0.0004	0.0762
LOC106783330	106783330	uncharacterized LOC106783330		-1.67	0.0001	0.0417
LOC111767520	111767520	uncharacterized LOC111767520		-1.60	0.0003	0.0669
LOC111768661	111768661	translation initiation factor IF-2-like		-1.19	0.0006	0.0897
LOC111768809	111768809	uncharacterized LOC111768809		-2.47	0.0000	0.0113
LOC111771758	111771758	GTPase IMAP family member 5-like		-0.92	0.0001	0.0474
LOC111774331	111774331	uncharacterized LOC111774331		-1.17	0.0003	0.0620
MCF2L	100067048	MCF.2 cell line derived transforming sequence like	*MCF2L*	-0.74	0.0001	0.0405
MICAL1	100066627	microtubule associated monooxygenase, calponin and LIM domain containing 1	*MICAL1*	-0.90	0.0000	0.0263
MMP25	100068942	matrix metallopeptidase 25	*MMP25*	-1.10	0.0003	0.0619
NAAA	100057831	N-acylethanolamine acid amidase	*NAAA*	-1.04	0.0003	0.0642
NUP210L	100056659	nucleoporin 210 like	*NUP210L*	-1.11	0.0002	0.0553

Gene symbol	Gene ID	Gene description	human Gene symbol	logFC	PValue	FDR
PDE10A	100050311	phosphodiesterase 10A	*PDE10A*	-2.44	0.0000	0.0195
PLCB2	100057315	phospholipase C beta 2	*PLCB2*	-0.93	0.0005	0.0861
PLXNA3	100058349	plexin A3	*PLXNA3*	-1.31	0.0005	0.0861
PREX1	100071328	phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 1	*PREX1*	-0.68	0.0006	0.0861
RAB44	100629655	RAB44, member RAS oncogene family	*RAB44*	-1.29	0.0005	0.0861
RYR1	100034090	ryanodine receptor 1	*RYR1*	-1.31	0.0004	0.0737
SLC8B1	100056481	solute carrier family 8 member B1	*SLC8B1*	-0.66	0.0003	0.0676
SNRNP70	100054907	small nuclear ribonucleoprotein U1 subunit 70	*SNRNP70*	-0.74	0.0002	0.0591
THBS2	100050044	thrombospondin 2	*THBS2*	-3.13	0.0000	0.0294
TIA1	100050503	TIA1 cytotoxic granule associated RNA binding protein	*TIA1*	-0.55	0.0006	0.0861
TNNT2	100146343	troponin T2, cardiac type	*TNNT2*	-2.20	0.0001	0.0417
TNXB	100059315	tenascin XB	*TNXB*	-1.19	0.0000	0.0240
TOP3B	100051153	DNA topoisomerase III beta	*TOP3B*	-0.92	0.0002	0.0602
TPCN2	102150167	two pore segment channel 2	*TPCN2*	-0.76	0.0002	0.0591
VGLL3	100069930	vestigial like family member 3	*VGLL3*	-1.41	0.0006	0.0883
WDR90	100066920	WD repeat domain 90	*WDR90*	-0.75	0.0005	0.0861
ZBP1	100055754	Z-DNA binding protein 1	*ZBP1*	-0.96	0.0000	0.0294
ZNF333	100064631	zinc finger protein 333	*ZNF333*	-0.67	0.0007	0.0998

3. 4 Overrepresented functional categories

The DEGs were analyzed for overrepresented functional categories and pathways in fertile or subfertile mares using the Metascape enrichment analysis tool (Table 4, Fig. 3, Supplementary data Table S2). The analyses were performed separately for genes upregulated or downregulated in subfertile mares compared to fertile mares, uploading the corresponding human NCBI gene IDs. Eighty-five genes of the downregulated DEGs and twelve of the upregulated DEGs could be assigned to a human gene symbol.

For genes with lower expression in subfertile compared to fertile mares, functional categories such as ‘extracellular matrix (ECM)’, ‘lymphocyte mediated immunity’, ‘immune response’, ‘positive regulation of cytosolic calcium ion concentration’ and ‘peptidyl-tyrosine phosphorylation’ were found as overrepresented. The most significantly enriched KEGG pathways were ‘ECM-receptor interaction’ (Fig. 4), ‘focal adhesion’ and ‘PI3K-Akt signaling pathway’. DEGs upregulated in subfertile mares were enriched for ‘monocarboxyl acid transmembrane transporter activity’ and ‘protein targeting’.

Table 4. Metascape enrichment analysis of DEGs subfertile vs fertile mares.

Most informative categories of Metascape enrichment analysis	LogP[1]	Genes[2]
Genes with lower expression in subfertile vs. fertile mares		
Extracellular matrix, ECM-receptor interaction, Focal adhesion, Collagen trimer, PI3-Akt signaling pathway	-7.8	*COL4A1,COL4A2,COL6A1,FN1,ITGB3,THBS2,TNXB,COL16A1,C1QA,C1QB,ACHE,MMP25,FGFR1,JAK3,SLC39A8,ADAMTS7,PNPLA2,ANO8,DLG4,TIA1,PLCB2,PLXNA3,LLGL1,ACKR3,PDE2A,RYR1,ERFE,AKR1C4,CLK2*
Lymphocyte mediated immunity, complement activation, adaptive immune response	-4.4	*C1QA,C1QB,HLAA,IGHA1,IGHG1,IGHM,CLCF1,ULBP3,JAK3,FN1,DLG4,ENPP3,STAC,ACKR3,LAT,PREX1,IRF8,ITGB3*
Positive regulation of cytosolic calcium ion concentration, muscle contraction	-3.8	*DLG4,PLCB2,RYR1,NMU,ACKR3,SLC8B1,TPCN2,SLC39A8,ERFE,COL6A1,STAC,ITGB3,PNPLA2,FGFR1,SYT12,ATP8B2,ANO8,TNNT2*
Glycosaminoglycan binding	-3.8	*COL16A1,FGFR1,FN1,IGHM,THBS2,TNXB,ITGB3,PREX1,IGHA1,SLC8B1*
Regulation of immune effector process	-3.7	*C1QA,C1QB,DNASE1L3,HLA-A,IGHG1,JAK3,ENPP3,CLCF1,ULBP3*
Peptidyl-tyrosine phosphorylation	-3.5	*DLG4,FGFR1,FN1,ITGB3,JAK3,LAT,CLK1,CLK2,CLCF1,IGHG1,COL16A1,PLCB2*
Inorganic anion transport	-3.5	*DLG4,FGFR1,FN1,ITGB3,JAK3,LAT,CLK1,CLK2,CLCF1,IGHG1,COL16A1,PLCB2*
Phosphoric diester hydrolase activity	-3.4	*PDE2A,ENPP3,PLCB2,PDE10A,ENTPD6*
Receptor internalization, receptor-mediated endocytosis	-3.1	*ACHE,DLG4,ITGB3,ACKR3,IGHA1*
Hallmark Myogenesis, calcium ion binding	-3.1	*ACHE,COL4A2,RYR1,TNNT2,HSPB8,ENTPD6,SYT12,C1QA,DLG4,DNASE1L3,ENPP3,PLCB2,THBS2,RAB44*
Genes with higher expression in subfertile vs. fertile mares		
Monocarboxylic acid transmembrane transport activity	-5.6	*SLC10A2, SLC16A5, SLC16A9*
Protein targeting	-3.0	*RPL21,RPS17, TOMM7*

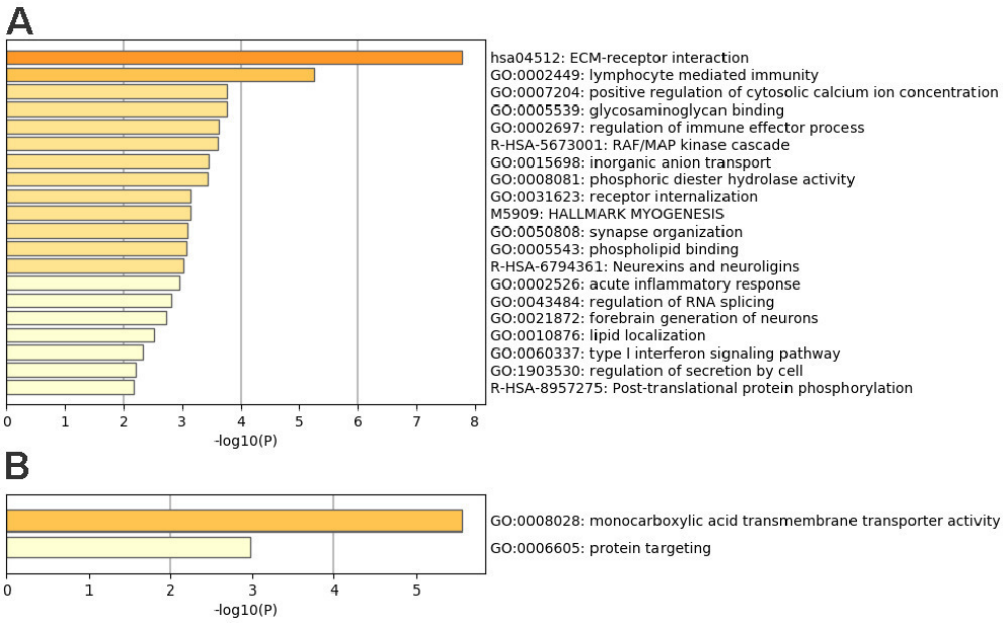

Figure 3. Metascape analysis: A) Top 20 most enriched terms in genes downregulated in subfertile mares compared to fertile mares. B) Enriched terms in genes upregulated in subfertile mares compared to fertile mares.

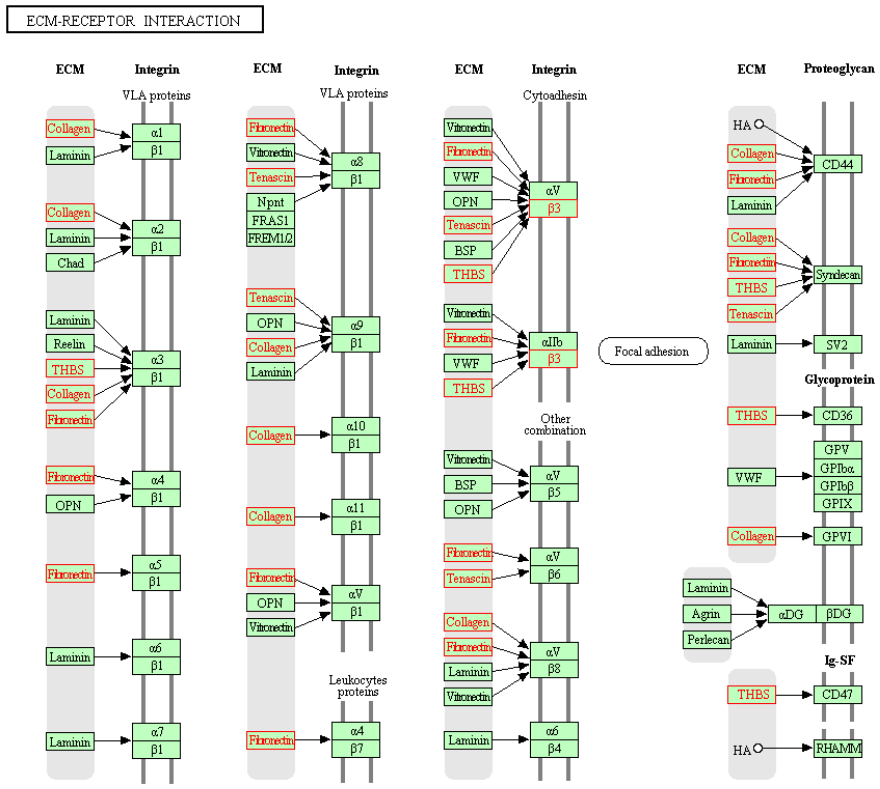

Figure 4. KEGG-Pathway ECM-receptor interaction: Genes belonging to ECM-receptor interaction are significantly downregulated in subfertile mares (red).

3.5 Validation of RNA-seq result by quantitative real-time RT-PCR

Expression differences found by RNA sequencing were confirmed by qRT-PCR for ten selected DEGs (Table 5). The qRT-PCR and RNA-seq relative expression values correlated well for the 22 analyzed samples (Fig. 5).

Table 5. Quantification of selected genes with quantitative real-time RT-PCR and comparison with RNA sequencing results.

	q-RT-PCR		RNA-Sequencing	
Gene symbol	logFC	p value	logFC	FDR
SLC10A2	1.7	0.0497	1.8	0.0368
SLC16A9	1.2	0.0127	1.0	0.0200
ITGB3	-1.9	0.0095	-1.3	0.0113
THBS2	-2.1	0.0269	-3.2	0.0080
FN1	-1.6	0.0156	-2.5	0.0065
COL6A1	-1.4	0.0313	-1.9	0.0460
ACKR3	-1.3	0.0084	-1.6	0.0242
PDE10A	-1.7	0.0130	-2.4	0.0080
MMP25	-0.7	0.0388	-1.1	0.0375
ENPP3	-1.2	0.0143	-2.2	0.0200

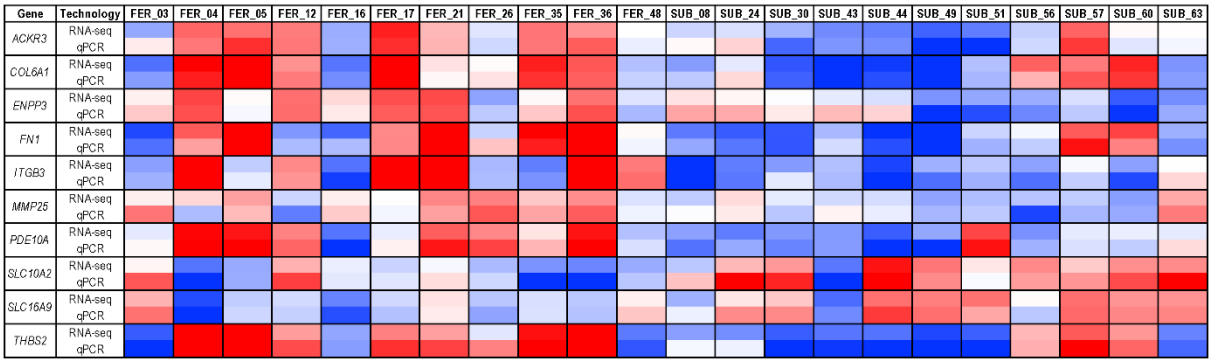

Figure 5. Confirmation of RNA sequencing results: Heatmap of quantitative real-time RT-PCR and RNA-seq data for ten selected genes. To illustrate correlation of RNA seq and qPCR data relative expression levels (mean-centered log2 expression values) are shown as a heatmap. Red color means higher and blue lower expression levels of the gene compared to the mean of all samples, respectively (from 2 over 0 to -2).

4. Discussion

To our knowledge, this is the first study investigating the relationship between uterine transcriptome and fertility in mares using cytobrush samples collected during estrus. Our study showed that sufficient amounts of high-quality RNA can be isolated from uterine cytobrush samples collected from mares. All obtained RNA samples showed RINs between 8.9 and 10 and revealed a minimum of 560 ng total RNA. In contrast to biopsy samples, the cytobrush technique does not provide information about gene expression of the whole endometrium as the cytobrush tends to collect only superficial parts of the endometrium and uterine fluid. To our knowledge, there is no study that examined, which material is exactly collected by the cytobrush. However, cytological examinations of cytobrush samples show primarily endometrial epithelial cells and infiltrating polymorphonuclear neutrophils (PMNs) (36). Comparing biopsy and cytobrush samples in cattle, stromal and endothelial cells were enriched in biopsy samples, while endometrial epithelial cell and immune cell markers were enriched in cytobrush samples (49). A previous study in mares at the time of recognition of

pregnancy showed that the strongest gene expression differences between pregnant and cyclic state are localized in the luminal epithelium (50). Therefore, we also expected the highest differences between fertile and subfertile mares in the endometrial epithelium, which is collected with the cytobrush samples.

The comparative transcriptome analysis of cytobrush samples collected during estrus revealed significant differences in the intrauterine gene expression between subfertile mares without clinical signs of reproductive diseases and normal fertile mares. Estrus was selected to allow easy sampling through the open cervix and to develop markers for the evaluation of fertility in mares before insemination based on routine cytobrush sampling. Early diagnosis of subfertile mares gives the possibility to improve the fertility of the mare with an optimized breeding management. In the present study, the mares were divided into fertile and subfertile mares according to pregnancy diagnosis after artificial inseminations during one breeding season. Mares becoming pregnant after only one AI were assumed fertile, mares that failed to conceive after at least three AIs were classified as subfertile. However, we are aware that probably not all mares classified as fertile are really fertile. Also, subfertile mares could become pregnant just by chance with the first AI and were considered as fertile in our classification. This could be also a reason why the hierarchical cluster analysis of the identified DEGs did not show a complete and clear separation of the two groups of mares into two clusters. Some mares showed intermediate expression patterns or patterns more similar to the other group. Classification after multiple AIs and pregnancy diagnosis, as in the study of Killeen et al. (51) in cattle, would have been better, but was not possible in the stud farm due to financial, logistical and ethical reasons. Moreover, we cannot exclude, if fertility was affected by the stallion, although we included only mares in our study inseminated with chilled semen from fertile stallions. Furthermore, in the subfertile mares, samples were collected after at least two unsuccessful inseminations in previous cycles, whereas in the fertile mares the samples were taken before the first insemination in the breeding season. Therefore, previous inseminations in the subfertile mares could have an influence on the intrauterine transcriptome.

In total, 114 genes were found as differentially expressed between subfertile and fertile mares. Quantitative real-time RT-PCR confirmed the results for 10 selected DEGs. The majority of the DEGs showed uniform expression in the respective fertility groups suggesting them as good candidates for RNA markers to diagnose subfertile mares without clinical signs of reproductive diseases and to predict fertility. A part of the DEGs (cluster 2 in the hierarchical cluster analysis) showed only increased expression in some of the samples of the fertile group, making these DEGs unsuitable as markers. No similarity could be found in these samples with respect to data records of these mares, which would distinguish them from the other samples. It seems likely that low mRNA levels of the genes of cluster 2 do not cause subfertility as a sole factor but may decrease fertility in conjunction with other disturbances.

According to the functional enrichment analysis, genes related to ‘extracellular matrix’, ‘ECM-receptor interaction’, ‘focal adhesion’, ‘immune response’, and ‘cytosolic calcium concentration’ may play an important role in fertility, as intrauterine gene expression levels related to these categories were lower in subfertile mares compared to fertile mares. Genes with higher mRNA concentrations in samples derived from subfertile mares were related to ‘monocarboxyl acid transmembrane transport activity’ and ‘protein targeting’. A comparison with human fertility studies is difficult as these studies were often not conducted at estrus like our study. In the following, selected DEGs potentially important for fertility are discussed.

ECM related genes downregulated in subfertile mares

Genes with lower mRNA levels in samples collected from subfertile mares were most significantly enriched for functional categories associated with ‘extracellular matrix (ECM)’,

'collagen trimer', 'extracellular matrix organization', 'ECM-receptor interaction', 'focal adhesion', and 'angiogenesis'. In our study the mRNA expression of 4 collagens (*COL4A1, COL4A2, COL6A1, COL16A1*), 4 ECM glycoproteins (*FN1*, *THBS2,* tenascin XB - *TNXB,* acetylcholinesterase (Cartwright blood group) - *ACHE*), the ECM receptor *ITGB3* and two ECM modifiers (*ADAMTS5, MMP25)* was lower in subfertile mares compared to fertile mares. The genes *ITGB3*, *THBS2*, *TNXB*, *COL4A1, COL6A1*, *COL4A2* and *FN1* are involved in 'ECM receptor interaction' and 'focal adhesion' pathway, referring to the interaction of cells with the ECM, whereby integrins or proteoglycans establish the link between ECM and cells.

Previous studies showed that ECM-related uterine gene expression varied during the equine estrous cycle (24, 25). Genes related to ECM, focal adhesion and angiogenesis were upregulated during estrus in equine endometrium (24, 25), which suggests an important function of these genes during estrus, that has not been clarified yet. Our study indicates that higher expression levels of genes related to ECM and focal adhesion during estrus might be important for fertility, as subfertile mares showed lower mRNA levels of these genes.

Genes related to ECM and cell adhesion were identified in fertility studies in humans at time of implantation as dysregulated genes in the endometrium of women with unexplained infertility or with recurrent miscarriage (27, 29, 52). For instance the gene *ITGB3*, one DEG of cluster 2, is a marker of endometrial receptivity at the window of implantation in women and mice (53-55). However, not much is known about the function of ECM related uterine genes and its relation to fertility at estrus, the time point of our sampling, as most human fertility studies are conducted during secretory phase.

ENPP3 protein expression was only found in cyclic and not in postmenopause endometria of women indicating a relation with fertility (56). In the study of Klein et al. (57), uterine *ENPP3* mRNA was upregulated in pregnant mares compared to cyclic mares during the time of maternal recognition of pregnancy 13.5 days after ovulation. Different DEGs are involved in ECM remodeling and collagen organization. Matrix metallopeptidase 25 (*MMP25*) and ADAM metallopeptidase with thrombospondin type 1 motif 7 (*ADAMTS7*), both involved in remodeling and breakdown of ECM, were downregulated in subfertile mares. MMP25 is able to clear the ECM components type IV-collagen, gelatin, fibronectin and fibrin *in vitro* (58). The gene *TNXB* encodes the extracellular matrix glycoprotein tenascin XB. Tenascin plays a role in the organization and formation of elastin and fibrillar collagen in the ECM and is important for tissue structure and elastic fiber stability (59). Tenascin-deficient women were reported to have a higher risk of complications during pregnancy such as preterm premature rupture of membranes (60, 61), but *Tnx* knockout mice showed no or only mild pregnancy-related abnormalities (61). A function of *TNXB* during estrus regarding fertility was not reported yet. Thrombospondin 2 modulates collagen fibrillogenesis and inhibits angiogenesis (62). Overall, dysregulations of ECM components involved in focal adhesion could be related to disturbances of endometrial remodeling during the estrous cycle affecting uterine interactions with the embryo and/or sperm.

Immune-related genes downregulated in subfertile mares

Metascape analysis revealed overrepresentation of DEGs related to 'lymphocyte mediated immunity', 'classical pathway of complement activation', 'adaptive immune response' and 'regulation of immune effector process' in genes downregulated in subfertile mares in comparison to fertile mares. At the time of sampling in midestrus, one to three days before insemination, there was no evidence of inflammation neither clinically nor facultative pathogens in the microbial culture or positive findings in the cytological examination in the selected mares for RNA-seq. Marth et al. (25) and Gebhardt et al. (24) described in previous

studies that genes associated with immune function are upregulated in the equine uterus during estrus in healthy mares. The authors suspected that the endometrium is preparing for mating during estrus and the required uterine clearance after mating with upregulation of immune-related genes and infiltration by immune cells, respectively. During estrus, the activity of the uterine immune system is increased due to intrauterine inoculation of foreign material such as bacteria, debris and sperm through the open cervix and by mating. A rapid high peak of inflammatory response after invasion of pathogens and a tight balance between pro- and anti-inflammatory factors are required to eliminate pathogens and avoid prolonged inflammation (23, 63). After AI an inflammatory response is important to effectively remove excessive spermatozoa, seminal plasma and other contaminants from the uterus before the embryo reaches the uterus 5-6 days after fertilization (7). For instance, the innate immune response of mares susceptible to PBIE is altered after breeding or bacterial inoculation, and characterized by abnormal imbalance in pro-inflammatory and anti-inflammatory cytokines (8, 20, 21, 23). In our study, the fertile mares with upregulated DEGs related to immune response might therefore be better prepared for these challenges during estrus than the subfertile mares. Even so clinically or in rectal ultrasound no differences between the fertile and subfertile mares could be seen in estrus or after insemination. No mares showed prolonged inflammation after insemination in rectal ultrasonic examinations (64). Also the recently identified diagnostic markers (*EBD1, LYZ* and *SLPI*) for susceptibility to PBIE were not differentially expressed between subfertile and fertile mares (22), supporting the assumption that subfertility in mares was not due to susceptibility to PBIE. Whereby these diagnostic markers were developed for uterine biopsy samples and not for cytobrush samples. The mRNAs encoding cytokines, which differ in susceptible mares after insemination, were not differentially expressed at estrus in the fertile and subfertile mares.

Different genes encoding proteins involved in activation of lymphocytes were downregulated in subfertile mares. The major histocompatibility complex, class I, A (*HLA-A)* and linker for activation of T cells (*LAT)* have a role in T cell activation (65). The genes *IGHG1* (*LOC102150085*, *LOC102149846*), *IGHA1* (*LOC106781303*), *IGHM* (*LOC100055483*), encoding heavy constants of immunoglobulins, and the DEGs cardiotrophin like cytokine factor 1 (*CLCF1)* and Janus kinase 3 (*JAK3)* are involved in B-cell activation (66, 67). Genes encoding immunoglobulins were also lower expressed in endometrial biopsies of subfertile heifers compared to high fertile heifers on day 14 after estrus (35). The *CLCF1* gene has been found as expressed in the equine adult ovary (67) and in equine endometrium on day 16 of pregnancy (68). Janus kinase 3 (JAK3) mediates signaling events in innate and adaptive immunity. Mutations in the *JAK3* gene leads to severe combined immunodeficiency in humans and mice due to a disturbed T- and B-lymphocyte development and function (69, 70). In innate immune cells, *JAK3* inhibition was reported to enhance the toll like receptor mediated production of pro-inflammatory cytokines while suppressing the anti-inflammatory cytokine IL-10 (71). Besides immune function, *in vitro* cultured bovine endometrial cell experiments showed that bovine endometrial cells treated with JAK3 had an increased cell viability, suggesting that JAK3/STAT pathway may also be involved in cell proliferation in endometrial cells (72).

Another downregulated gene of particular interest is the cytomegalovirus UL 16 binding protein 1 gene (*LOC100063831*), one DEG from cluster 2. The human orthologue *ULBP3* gene encodes a ligand of the immunoreceptor NKG2D, which is mainly expressed by natural killer cells (NK) and T cells (73). As a ligand of NKG2D of natural NK cells, ULBP stimulates cytokine and chemokine production from NK cells. Furthermore, decidual NK cells play a role in early human and mouse pregnancy by regulation trophoblast invasion and vascularization in the decidua (74-76).

Genes involved in the classical pathway of complement activation were downregulated in uterine cytobrush samples of subfertile mares. The genes *C1QA* and *C1QB* encode the A- and B chain polypeptide of complement component C1q, which is the first component of the classical pathway of complement activation and is activated through binding of immunoglobulins (77). Genes encoding immunoglobulins were also downregulated in subfertile mares. Besides the function in the complement pathway, C1q also possesses a physiological role in trophoblast invasion, spiral arteries remodeling and placentation in the mouse (78-80). C1q-deficient mice showed pregnancy disorders characterized by increased fetal death and signs similar to human pre-eclampsia.

Different genes participating in regulation of innate immune cells showed lower uterine mRNA levels in subfertile mares compared to fertile mares. For instance, phosphatidylinositol-3,4,5-triphosphate dependent Rac exchange factor 1 (*Prex1)* is strongly expressed in leukocytes and is involved in neutrophil recruitment during inflammation (81). Besides the function in extracellular matrix remodeling, MMPs also participate in regulation of inflammation and innate immunity for example by activating and degrading chemotactic molecules (82). Previous studies showed that the expression of several MMPs (*MMP1, MMP2, MMP3, MMP7, MMP8, MMP9*) was upregulated in the endometrium of mares after inoculation of *Streptococcus zooepidemicus* or *E. coli* or after insemination (63, 83). *MMP25*, which was differentially expressed between subfertile and fertile mares in our study, has not been described previously in the equine uterus. In mice, *Mmp25* has been found predominantly in peripheral blood leukocytes. It participates in the regulation of innate immune response. *Mmp25*-deficient mice are fertile but show a deficiency in their innate immune response (84). The downregulated *ACKR3*, also known as *CXCR7* or decoy receptor D6, regulates innate and adaptive immunity with his activity as a decoy and scavenger receptor for inflammatory chemokines and thus is involved in the control of inflammation (85). In mares, *ACKR3* is expressed significantly higher in endometrium during estrus compared to diestrus (25) and was increased after inoculation of *E. coli* (63). *ACKR3* might play a role also in the balancing between protection of the developing embryo and tolerance of its hemiallogeneic tissues (85). *Ackr3*-deficient mice showed an earlier and exacerbated inflammatory response in a model of skin inflammation, with high levels of inflammatory chemokines (86). Although *Ackr3*-deficient mice were fertile, exposure to LPS or antiphospholipid autoantibodies resulted in higher levels of inflammatory CC chemokines and increased leukocyte infiltrate in placenta, causing an increased rate of fetal loss (87). The higher level of *ACKR3* mRNA expression in fertile mares could affect fertility by preventing an excessive inflammatory response to insemination and might also be important in establishment of pregnancy.

Cysteine rich secretory protein 2 mRNA (*CRISP2*) was not strongly represented in the uterus samples; nevertheless, a significantly lower level of *CRISP2* mRNA was observed in subfertile compared to fertile mares. *CRISP2* plays an important role in sperm motility, sperm capacitation and sperm-egg fusion and is positively correlated to male fertility in humans (88), mice (89), and bulls (90). In stallions only *CRISP3* correlates with stallion fertility (91). Equine *CRISP3*, highly presented in seminal plasma, suppresses binding between PMNs and viable spermatozoa in the reproductive tract of the mare (92). Cysteine rich secretory proteins expression and proteins were also found in some studies in the female reproductive tract of mice (93) and humans (94). *In vitro* experiments showed that epithelial and neutrophil-derived *CRISP3* plays a role in mouse postmenstrual endometrial repair and regeneration and *CRISP3* increases adhesion and proliferation of human epithelial cells (94). To our knowledge, it is the first time that *CRSIP2* was described in the equine endometrium. In a recently published study, Klein et al. (68) reported that *CRISP3* is expressed in conceptus tissue at day 16 of pregnancy.

Taken together, genes for activation of lymphocytes and genes involved in regulation of immune cells seem to play an important role in equine uterine receptivity.

Cytosolic calcium ion concentration related genes downregulated in subfertile mares

Furthermore, DEGs were found to be involved in 'positive regulation of cytosolic calcium ion concentration' and 'muscle contraction'. Calcium ion mobilization in cells occurs via Ca^{2+} channels located in the plasma membrane or via intracellular release of calcium from sarcoplasmic reticulum or acidic lysosomal stores (95). For instance, the DEG *SLC8B1* encodes a mitochondrial Na^{+}/Ca^{2+} exchanger (NCLX) (96). The DEG two pore calcium channel 2 (*TPC2)* encodes two pore segment channel 2 protein, a key component of the NAADP receptor, necessary for NAADP-mediated Ca^{2+} release from lysosome related organelles (95). Intracellular Ca^{2+} release from the sarcoplasmic reticulum occurs via ryanodine receptor, encoded by the DEG *RYR1*, or inositol 1,4,5-triphospate (IP3)-receptor. Through activation of G protein linked receptor Phospholipase C ß (*PLCB2*) it hydrolyzes phosphatidylinositol 4,5-biphosphate into diacylglycerol (DAG) and IP3. Increased concentration of IP3 leads to Ca^{2+} secretion from the sarcoplasmic reticulum (97). For example, neuromedin U, encoded by the DEG *NMU*, binds to the G-protein coupled NMU2-receptor, which activates Phospholipase C and has an uterocontractile effect (98). In accordance with our study, another member of the neuromedin family neuromedin B was upregulated in endometrial biopsy samples of fertile cows compared to infertile ones (32). The cytosolic calcium concentration controls processes such as metabolism, secretion, fertilization, proliferation, and smooth muscle contraction (99). Uterine contractility in mares after insemination is important to carry sperm toward the oviduct and to eliminate excessive sperm and contaminants from uterus (100, 101).

Genes upregulated in subfertile mares

Genes upregulated in subfertile mares without clinical signs of uterine diseases compared to fertile mares were mainly related to 'monocarboxylic acid transmembrane transporter activity' (*SLC10A2, SLC16A5, SLC16A9*) and protein targeting (*RPS17, RPL21, TOMM7, 60S ribosomal protein L26-like pseudogene, 40S ribosomal protein S2 pseudogene*). Several genes encoding ribosomal proteins were also more abundant in low-fertility heifers compared to high fertility heifers at day 14 post estrus (35).

The gene *SLC10A2* encodes a sodium dependent bile acid transporter highly expressed in the liver and intestine (102). Higher abundance of *SLC10A2* were observed in bovine follicular fluid of lactating cows than in heifers suggesting that increased bile acids within follicular microenvironment may be the reason for fertility problems in lactating cows (103). The DEGs *SLC16A5* and *SLC16A9* encode monocarboxylate transporters MCT6 and MCT9. MCT6, highly expressed in placenta and kidney, transports bumetanide, nateglinide, probenecid and PFG2α. The transport of bumetanide is pH dependent and elevated by extracellular acidic pH (104, 105). MCT9 is involved in transport of pyruvic acid, lactic acid and carnitine (106). The upregulated crystallin lambda 1 gene (*CRYL1)* is involved in ß oxidation of fatty acid and the upregulated ELOVL fatty acid elongase 2 (*ELOVL2)* in elongation of very long fatty acids such as arachidonic acid. Arachidonic acid in turn is a precursor of prostaglandins, which means that ELOVL2 has an influence in prostaglandin metabolism (107, 108). In our study, higher abundance of uterine *ELOVL2* seems to have a negative effect on fertility in mares.

5. Conclusions

In conclusion, cytobrush samples can be used for uterine transcriptome analysis in the mare. Our study revealed significant differences in the uterine transcriptome at estrus between fertile and subfertile mares without clinical signs of uterine diseases. We have identified a large number of DEGs, which are potential candidates for RNA biomarkers for the prognosis of subfertility in the mare. In further studies, these results have to be validated in a higher number of mares.

Abbreviations

AI	Artificial insemination
CPM	Counts per million
DEG	Differentially expressed gene
DNA	Deoxyribonucleic acid
cDNA	complementary DNA
ECM	Extracellular matrix
FB	First Breeders
FB-P	First Breeder- pregnant
FB-N	First Breeder- not pregnant
FDR	False discovery rate
i.v.	intravenous
KEGG	Kyoto Encyclopedia of Genes and Genomes
logFC	log2-fold-change
mRNA	Messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
NAADP	Nicotinic acid adenine dinucleotide phosphate
NKG2D	Natural Killer Group 2D
PBIE	Persistent breeding induced endometritis
RB	Repeat Breeders
RB-P	Repeat Breeder- pregnant
RB-N	Repeat Breeder- not pregnant
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RT-PCR	Reverse transcription- Polymerase Chain Reaction
qRT-PCR	Quantitative RT-PCR

References

1. Riddle WT, LeBlanc MM, Stromberg AJ. Relationships between uterine culture, cytology and pregnancy rates in a Thoroughbred practice. Theriogenology. 2007;68(3):395-402.
2. Waelchli RO. Endometrial biopsy in mares under nonuniform breeding management conditions: Prognostic value and relationship with age. Can Vet J. 1990;31(5):379-84.
3. Kareskoski M, Venhoranta H, Virtala AM, Katila T. Analysis of factors affecting the pregnancy rate of mares after inseminations with cooled transported stallion semen. Theriogenology. 2019;127:7-14.
4. Morris LHA, Allen WR. Reproductive efficiency of intensively managed Thoroughbred mares in Newmarket. Equine Vet J. 2002;34(1):51-60.
5. Traub-Dargatz JL, Salman MD, Voss JL. Medical problems of adult horses, as ranked by equine practitioners. J Am Vet Med Assoc. 1981;198(10):1745–7.
6. LeBlanc MM. Advances in the diagnosis and treatment of chronic infectious and post-mating-induced endometritis in the mare. Reprod Domest Anim. 2010;45 Suppl 2:21-7.
7. Christoffersen M, Troedsson M. Inflammation and fertility in the mare. Reprod Domest Anim. 2017;52 Suppl 3:14-20.
8. Woodward EM, Christoffersen M, Campos J, Betancourt A, Horohov D, Scoggin KE, et al. Endometrial inflammatory markers of the early immune response in mares susceptible or resistant to persistent breeding-induced endometritis. Reproduction. 2013;145(3):289-96.
9. LeBlanc MM, Causey RC. Clinical and subclinical endometritis in the mare: both threats to fertility. Reprod Domest Anim. 2009;44 Suppl 3:10-22.
10. Nielsen JM. Endometritis in the mare: a diagnostic study comparing cultures from swab and biopsy. Theriogenology. 2005;64(3):510-8.
11. LeBlanc MM, Magsig J, Stromberg AJ. Use of a low-volume uterine flush for diagnosing endometritis in chronically infertile mares. Theriogenology. 2007;68(3):403-12.
12. Christoffersen M, Brandis L, Samuelsson J, Bojesen AM, Troedsson MH, Petersen MR. Diagnostic double-guarded low-volume uterine lavage in mares. Theriogenology. 2015;83(2):222-7.
13. Buczkowska J, Kozdrowski R, Nowak M, Ras A, Staroniewicz Z, Siemieniuch MJ. Comparison of the biopsy and cytobrush techniques for diagnosis of subclinical endometritis in mares. Reprod Biol Endocrinol. 2014;12:27.
14. Kenney RM, Doig PA. Equine endometrial biopsy. In: Current Therapy in Theriogenology, 2nd ed; Morrow, DA, Saunders, WB, Philadelphia. 1986:723-9.
15. Overbeck W, Witte TS, Heuwieser W. Comparison of three diagnostic methods to identify subclinical endometritis in mares. Theriogenology. 2011;75(7):1311-8.
16. Singh J, Behal A, Singla N, Joshi A, Birbian N, Singh S, et al. Metagenomics: Concept, methodology, ecological inference and recent advances. Biotechnol J. 2009;4(4):480-94.
17. Ferris RA. Endometritis: Diagnostic Tools for Infectious Endometritis. Vet Clin North Am Equine Pract. 2016;32(3):481-98.
18. Rhoads DD, Cox SB, Rees EJ, Sun Y, Wolcott RD. Clinical identification of bacteria in human chronic wound infections: culturing vs. 16S ribosomal DNA sequencing. BMC Infect Dis. 2012;12(321).
19. Nielsen JM, Troedsson MH, Pedersen MR, Bojesen AM, Lehn-Jensen H, Zent WW. Diagnosis of endometritis in the mare based on bacteriological and cytological examinations of the endometrium: Comparison of results obtained by swabs and biopsies. J Equine Vet Sci. 2010;30(1):27-30.
20. Fumuso E, Giguère S, Wade J, Rogan D, Videla-Dorna I, Bowden RA. Endometrial IL-1β, IL-6 and TNF-α, mRNA expression in mares resistant or susceptible to post-breeding endometritis. Vet Immunol Immunopathol. 2003;96(1-2):31-41.
21. Fumuso E, Aguilar J, Giguère S, David O, Wade J, Rogan D. Interleukin-8 (IL-8) and 10 (IL-10) mRNA transcriptions in the endometrium of normal mares and mares susceptible to persistent post-breeding endometritis. Effects of estrous cycle, artificial insemination and immunomodulation. Anim Reprod Sci. 2006;94(1-4):282-5.

22. Marth CD, Firestone SM, Hanlon D, Glenton LY, Browning GF, Young ND, et al. Innate immune genes in persistent mating-induced endometritis in horses. Reprod Fertil Dev. 2018;30(3):533-45.
23. Christoffersen M, Woodward E, Bojesen AM, Jacobsen S, Petersen MR, Troedsson MH, et al. Inflammatory responses to induced infectious endometritis in mares resistant or susceptible to persistent endometritis. BMC Vet Res. 2012;8:41.
24. Gebhardt S, Merkl M, Herbach N, Wanke R, Handler J, Bauersachs S. Exploration of global gene expression changes during the estrous cycle in equine endometrium. Biol Reprod. 2012;87(6):136.
25. Marth CD, Young ND, Glenton LY, Noden DM, Browning GF, Krekeler N. Effect of ovarian hormones on the healthy equine uterus: a global gene expression analysis. Reprod Fertil Dev. 2015.
26. Bersinger NA, Wunder DM, Birkhäuser MH, Mueller MD. Gene expression in cultured endometrium from women with different outcomes following IVF. Mol Hum Reprod. 2008;14(8):475-84.
27. Othman R, Omar MH, Shan LP, Shafiee MN, Jamal R, Mokhtar NM. Microarray profiling of secretoryphase endometrium from patients with recurrent miscarriage. Reprod Biol. 2012;12(2):183-99.
28. Tapia A, Gangi LM, Zegers-Hochschild F, Balmaceda J, Pommer R, Trejo L, et al. Differences in the endometrial transcript profile during the receptive period between women who were refractory to implantation and those who achieved pregnancy. Hum Reprod. 2008;23(2):340-51.
29. Altmäe S, Martínez-Conejero JA, Salumets A, Simón C, Horcajadas JA, Stavreus-Evers A. Endometrial gene expression analysis at the time of embryo implantation in women with unexplained infertility. Mol Hum Reprod. 2010;16(3):178-87.
30. Koler M, Achache H, Tsafrir A, Smith Y, Revel A, Reich R. Disrupted gene pattern in patients with repeated in vitro fertilization (IVF) failure. Hum Reprod. 2009;24(10):2541-8.
31. Wagener K, Pothmann H, Prunner I, Peter S, Erber R, Aurich C, et al. Endometrial mRNA expression of selected pro-inflammatory factors and mucins in repeat breeder cows with and without subclinical endometritis. Theriogenology. 2017;90:237-44.
32. Moran B, Butler ST, Moore SG, MacHugh DE, Creevey CJ. Differential gene expression in the endometrium reveals cytoskeletal and immunological genes in lactating dairy cows genetically divergent for fertility traits. Reprod Fertil Dev. 2017;29(2):274-82.
33. Moraes JGN, Behura SK, Geary TW, Hansen PJ, Neibergs HL, Spencer TE. Uterine influences on conceptus development in fertility-classified animals. Proc Natl Acad Sci U S A. 2018;115(8):E1749-E58.
34. Moore SG, Pryce JE, Hayes BJ, Chamberlain AJ, Kemper KE, Berry DP, et al. Differentially Expressed Genes in Endometrium and Corpus Luteum of Holstein Cows Selected for High and Low Fertility Are Enriched for Sequence Variants Associated with Fertility. Biol Reprod. 2016;94(1):19.
35. Geary TW, Burns GW, Moraes JGN, Moss JI, Denicol AC, Dobbs KB, et al. Identification of Beef Heifers with Superior Uterine Capacity for Pregnancy. Biol Reprod. 2016;95(2):47.
36. Ghasemi F, Gonzalez-Cano P, Griebel PJ, Palmer C. Proinflammatory cytokine gene expression in endometrial cytobrush samples harvested from cows with and without subclinical endometritis. Theriogenology. 2012;78(7):1538-47.
37. Diel de Amorim M, Gartley CJ, Foster RA, Hill A, Scholtz EL, Hayes A, et al. Comparison of Clinical Signs, Endometrial Culture, Endometrial Cytology, Uterine Low-Volume Lavage, and Uterine Biopsy and Combinations in the Diagnosis of Equine Endometritis. J Equine Vet Sci. 2016;44:54-61.
38. Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A, et al. Manipulation of FASTQ data with Galaxy. Bioinformatics. 2010;26(14):1783-5.
39. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: a platform for interactive large-scale genome analysis. Genome Res. 2005;15(10):1451-5.
40. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357-60.
41. Chen YS, Lun ATL, Smyth GK. Differential Expression Analysis of Complex RNA-seq Experiments Using edgeR. Front Probab Stat Sc. 2014:51-74.

42. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.
43. Bick JT, Zeng S, Robinson MD, Ulbrich SE, Bauersachs S. Mammalian Annotation Database for improved annotation and functional classification of Omics datasets from less well-annotated organisms. Database (Oxford). 2019;2019.
44. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, et al. TM4: A free, open-source system for microarray data management and analysis. Biotechniques. 2003;34(2):374-8.
45. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.
46. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14(2):178-92.
47. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 2012;13(134).
48. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8.
49. Cardoso B, Oliveira ML, Pugliesi G, Batista E, Binelli M. Cytobrush: A tool for sequential evaluation of gene expression in bovine endometrium. Reprod Domest Anim. 2017;52(6):1153-7.
50. Scaravaggi I, Borel N, Romer R, Imboden I, Ulbrich SE, Zeng S, et al. Cell type-specific endometrial transcriptome changes during initial recognition of pregnancy in the mare. Reprod Fertil Dev. 2019;31(3):496-508.
51. Killeen AP, Morris DG, Kenny DA, Mullen MP, Diskin MG, Waters SM. Global gene expression in endometrium of high and low fertility heifers during the mid-luteal phase of the estrous cycle. BMC Genomics. 2014;15(234).
52. Iwahashi M, Muragaki Y, Ooshima A, Yamoto M, Nakano R. Alterations in distribution and composition of the extracellular matrix during decidualization of the human endometrium. J Reprod Fertil. 1996;108:147-55.
53. Lessey BA, Castelbaum AJ, Sawin SW, Sun J. Integrins as markers of uterine receptivity in women with primary unexplained infertility. Fertil Steril. 1995;63(3):535-42.
54. Illera MJ, Cullinan E, Gui Y, Yuan L, Beyler SA, Lessey BA. Blockade of the alpha(v)beta(3) integrin adversely affects implantation in the mouse. Biol Reprod. 2000;62(5):1285-90.
55. Germeyer A, Savaris RF, Jauckus J, Lessey B. Endometrial beta3 integrin profile reflects endometrial receptivity defects in women with unexplained recurrent pregnancy loss. Reprod Biol Endocrinol. 2014;12:53.
56. Aliagas E, Vidal A, Torrejón-Escribano B, Taco Mdel R, Ponce J, de Aranda IG, et al. Ecto-nucleotidases distribution in human cyclic and postmenopausic endometrium. Purinergic Signal. 2013;9(2):227-37.
57. Klein C, Scoggin KE, Ealy AD, Troedsson MH. Transcriptional profiling of equine endometrium during the time of maternal recognition of pregnancy. Biol Reprod. 2010;83(1):102-13.
58. English WR, Velasco G, Stracke JO, Knäuper V, Murphy G. Catalytic activities of membrane-type 6 matrix metalloproteinase (MMP25). FEBS Lett. 2001;491:137-42.
59. Petersen JW, Douglas JY. Tenascin-X, collagen, and Ehlers-Danlos syndrome: tenascin-X gene defects can protect against adverse cardiovascular events. Med Hypotheses. 2013;81(3):443-7.
60. Anum EA, Hill LD, Pandya A, Strauss JF, 3rd. Connective tissue and related disorders and preterm birth: clues to genes contributing to prematurity. Placenta. 2009;30(3):207-15.
61. Egging DF, van Vlijmen-Willems I, Choi J, Peeters AC, van Rens D, Veit G, et al. Analysis of obstetric complications and uterine connective tissue in tenascin-X-deficient humans and mice. Cell Tissue Res. 2008;332(3):523-32.
62. Bornstein P, Kyriakides TR, Yang Z, Armstrong LC, Birk DE. Thrombospondin 2 modulates collagen fibrillogenesis and angiogenesis. J Investig Dermatol Symp Proc. 2000;5(1):61-6.

63. Marth CD, Young ND, Glenton LY, Noden DM, Browning GF, Krekeler N. Deep sequencing of the uterine immune response to bacteria during the equine oestrous cycle. BMC Genomics. 2015;16:934.
64. Canisso IF, Stewart J, Coutinho da Silva MA. Endometritis: Managing Persistent Post-Breeding Endometritis. Vet Clin North Am Equine Pract. 2016;32(3):465-80.
65. Zhang W, Sommers CL, Burshtyn DN, Stebbins CC, DeJarnette JB, Trible RP, et al. Essential Role of LAT in T Cell Development. Immunity. 1999;10:323-32.
66. Senaldi G, Stolina M, Guo J, Faggioni R, McCabe S, Kaufman SA, et al. Regulatory effects of novel neurotrophin-1/b cell-stimulating factor-3 (cardiotrophin-like cytokine) on B cell function. J Immunol. 2002;168(11):5690-8.
67. Hall SE, Upton RMO, McLaughlin EA, Sutherland JM. Phosphoinositide 3-kinase/protein kinase B (PI3K/AKT) and Janus kinase/signal transducer and activator of transcription (JAK/STAT) follicular signalling is conserved in the mare ovary. Reprod Fertil Dev. 2018;30(4):624-33.
68. Klein C. Novel equine conceptus?endometrial interactions on Day 16 of pregnancy based on RNA sequencing. Reprod Fertil Dev. 2015.
69. Macchi P, Villa A, Giliani S, Sacco MG, Frattini A, Porta F, et al. Mutations of Jak-3 gene in patients with autosomal severe combined immune deficiency (SCID). Nature. 1995;377(6544):65-8.
70. Thomis DC, Gurniak CB, Tivol E, Sharpe AH, Berg LJ. Defects in B lymphocyte maturation and T lymphocyte activation in mice lacking Jak3. Science. 1995;270:794-7.
71. Wang H, Brown J, Gao S, Liang S, Jotwani R, Zhou H, et al. The role of JAK-3 in regulating TLR-mediated inflammatory cytokine production in innate immune cells. J Immunol. 2013;191(3):1164-74.
72. Ndiaye K, Castonguay A, Benoit G, Silversides DW, Lussier JG. Differential regulation of Janus kinase 3 (JAK3) in bovine preovulatory follicles and identification of JAK3 interacting proteins in granulosa cells. J Ovarian Res. 2016;9:71.
73. Cosman D, Müllberg J, Sutherland C, Chin W, Armitage R, Fanslow W, et al. ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. Immunity. 2001;14:123-33.
74. Xie X, He H, Colonna M, Seya T, Takai T, Croy BA. Pathways participating in activation of mouse uterine natural killer cells during pregnancy. Biol Reprod. 2005;73(3):510-8.
75. Marlin R, Duriez M, Berkane N, de Truchis C, Madec Y, Rey-Cuille MA, et al. Dynamic shift from CD85j/ILT-2 to NKG2D NK receptor expression pattern on human decidual NK during the first trimester of pregnancy. PLoS One. 2012;7(1):e30017.
76. Hanna J, Goldman-Wohl D, Hamani Y, Avraham I, Greenfield C, Natanson-Yaron S, et al. Decidual NK cells regulate key developmental processes at the human fetal-maternal interface. Nat Med. 2006;12(9):1065-74.
77. Kishore U, Reid KBM. C1q:structure, function, and receptors. Immunopharmacology. 2000;49:159-70.
78. Singh J, Ahmed A, Girardi G. Role of complement component C1q in the onset of preeclampsia in mice. Hypertension. 2011;58(4):716-24.
79. Agostinis C, Bulla R, Tripodo C, Gismondi A, Stabile H, Bossi F, et al. An alternative role of C1q in cell migration and tissue remodeling: contribution to trophoblast invasion and placental development. J Immunol. 2010;185(7):4420-9.
80. Agostinis C, Tedesco F, Bulla R. Alternative functions of the complement protein C1q at embryo implantation site. J Reprod Immunol. 2017;119:74-80.
81. Dong X, Mo Z, Bokoch G, Guo C, Li Z, Wu D. P-Rex1 is a primary Rac2 guanine nucleotide exchange factor in mouse neutrophils. Curr Biol. 2005;15(20):1874-9.
82. Parks WC, Wilson CL, López-Boado YS. Matrix metalloproteinases as modulators of inflammation and innate immunity. Nat Rev Immunol. 2004;4(8):617-29.
83. Oddsdóttir C, Riley SC, Leask R, Edwards DR, Watson ED. Activities of matrix metalloproteinases-9 and -2 in uterine fluid during induced equine endometritis. Pferdeheilkunde. 2008;24:70-3.

84. Soria-Valles C, Gutierrez-Fernandez A, Osorio FG, Carrero D, Ferrando AA, Colado E, et al. MMP-25 Metalloprotease Regulates Innate Immune Response through NF-kappaB Signaling. J Immunol. 2016;197(1):296-302.
85. Borroni EM, Bonecchi R, Buracchi C, Savino B, Mantovani A, Locati M. Chemokine decoy receptors: new players in reproductive immunology. Immunol Invest. 2008;37(5):483-97.
86. Martinez de la Torre Y, Locati M, Buracchi C, Dupor J, Cook DN, Bonecchi R, et al. Increased inflammation in mice deficient for the chemokine decoy receptor D6. Eur J Immunol. 2005;35(5):1342-6.
87. Martinez de la Torre Y, Buracchi C, Borroni EM, Dupor J, Bonecchi R, Nebuloni M, et al. Protection against inflammation- and autoantibody- caused fetal loss by the chemokine decoy receptor D6. PNAS. 2007;104:2319-24.
88. Zhou JH, Zhou QZ, Lyu XM, Zhu T, Chen ZJ, Chen MK, et al. The expression of cysteine-rich secretory protein 2 (CRISP2) and its specific regulator miR-27b in the spermatozoa of patients with asthenozoospermia. Biol Reprod. 2015;92(1):28.
89. Lim S, Kierzek M, O'Connor AE, Brenker C, Merriner DJ, Okuda H, et al. CRISP2 Is a Regulator of Multiple Aspects of Sperm Function and Male Fertility. Endocrinology. 2019;160(4):915-24.
90. Arangasamy A, Kasimanickam VR, DeJarnette JM, Kasimanickam RK. Association of CRISP2, CCT8, PEBP1 mRNA abundance in sperm and sire conception rate in Holstein bulls. Theriogenology. 2011;76(3):570-7.
91. Novak S, Smith TA, Paradis F, Burwash L, Dyck MK, Foxcroft GR, et al. Biomarkers of in vivo fertility in sperm and seminal plasma of fertile stallions. Theriogenology. 2010;74(6):956-67.
92. Doty A, Buhi WC, Benson S, Scoggin KE, Pozor M, Macpherson M, et al. Equine CRISP3 modulates interaction between spermatozoa and polymorphonuclear neutrophils. Biol Reprod. 2011;85(1):157-64.
93. Reddy T, Gibbs GM, Merriner DJ, Kerr JB, O'Bryan MK. Cysteine-rich secretory proteins are not exclusively expressed in the male reproductive tract. Dev Dyn. 2008;237(11):3313-23.
94. Evans J, D'Sylva R, Volpert M, Jamsai D, Merriner DJ, Nie G, et al. Endometrial CRISP3 is regulated throughout the mouse estrous and human menstrual cycle and facilitates adhesion and proliferation of endometrial epithelial cells. Biol Reprod. 2015;92(4):99.
95. Tugba Durlu-Kandilci N, Ruas M, Chuang KT, Brading A, Parrington J, Galione A. TPC2 proteins mediate nicotinic acid adenine dinucleotide phosphate (NAADP)- and agonist-evoked contractions of smooth muscle. J Biol Chem. 2010;285(32):24925-32.
96. Khananshvili D. The SLC8 gene family of sodium-calcium exchangers (NCX) - structure, function, and regulation in health and disease. Mol Aspects Med. 2013;34(2-3):220-35.
97. Rebecchi MJ, Pentyala SN. Structure, Function, and Control of Phosphoinositide-Specific Phospholipase C. Physiol Rev. 2000;80(4):1291-335.
98. Nadeau-Vallée M, Boudreault A, Leimert K, Hou X, Obari D, Madaan A, et al. Uterotonic Neuromedin U Receptor 2 and Its Ligands Are Upregulated by Inflammation in Mice and Humans, and Elicit Preterm Birth. Biol Reprod. 2016;95(3):72.
99. Berridge MJ. The Inositol Trisphosphate/Calcium Signaling Pathway in Health and Disease. Physiol Rev. 2016;96(4):1261-96.
100. Katila T. Sperm-uterine interactions: a review. Anim Reprod Sci. 2001;68:267-72.
101. Troedsson MHT, Liu IKM, Crabo BG. Sperm transport and survival in the mare: a review. Theriogenology. 1998;50:807-18.
102. Xiao L, Pan G. An important intestinal transporter that regulates the enterohepatic circulation of bile acids and cholesterol homeostasis: The apical sodium-dependent bile acid transporter (SLC10A2/ASBT). Clin Res Hepatol Gastroenterol. 2017;41(5):509-15.
103. Sanchez R, Schuermann Y, Gagnon-Duval L, Baldassarre H, Murphy BD, Gevry N, et al. Differential abundance of IGF1, bile acids, and the genes involved in their signaling in the dominant follicle microenvironment of lactating cows and nulliparous heifers. Theriogenology. 2014;81(6):771-9.

104. Jones RS, Parker MD, Morris ME. Monocarboxylate Transporter 6-Mediated Interactions with Prostaglandin F2alpha: In Vitro and In Vivo Evidence Utilizing a Knockout Mouse Model. Pharmaceutics. 2020;12(3):201.
105. Murakami Y, Kohyama N, Kobayashi Y, Ohbayashi M, Ohtani H, Sawada Y, et al. Functional characterization of human monocarboxylate transporter 6 (SLC16A5). Drug Metab Dispos. 2005;33(12):1845-51.
106. Vadakedath S, Kandi V. Probable Potential Role of Urate Transporter Genes in the Development of Metabolic Disorders. Cureus. 2018;10(3):e2382.
107. González RS, Rodriguez-Cruz M, Maldonado J, Saavedra FJ. Role of maternal tissue in the synthesis of polyunsaturated fatty acids in response to a lipid-deficient diet during pregnancy and lactation in rats. Gene. 2014;549(1):7-23.
108. Kemiläinen H, Adam M, Mäki-Jouppila J, Damdimopoulou P, Damdimopoulos AE, Kere J, et al. The Hydroxysteroid (17beta) Dehydrogenase Family Gene HSD17B12 Is Involved in the Prostaglandin Synthesis Pathway, the Ovarian Function, and Regulation of Fertility. Endocrinology. 2016;157(10):3719-30.

Supplemental Tables

Supplemental Table 1

Equine EntrezGene ID	Gene symbol	Gene description	type of gene	Human EntrezGene ID	Gene symbol	Gene description	log_2 FC SUB/FER	P-Value	FDR
100072970	ACAP1	ArfGAP with coiled-coil, ankyrin repeat and PH domains 1	protein-coding	9744	ACAP1	ArfGAP with coiled-coil, ankyrin repeat and PH domains 1	-1.01	0.0000	0.0291
100069149	ACHE	acetylcholinesterase (Cartwright blood group)	protein-coding	43	ACHE	acetylcholinesterase (Cartwright blood group)	-1.68	0.0007	0.0998
100057501	ACKR3	atypical chemokine receptor 3	protein-coding	57007	ACKR3	atypical chemokine receptor 3	-1.60	0.0001	0.0474
100059959	ADAMTS7	ADAM metallopeptidase with thrombospondin type 1 motif 7	protein-coding	11173	ADAMTS7	ADAM metallopeptidase with thrombospondin type 1 motif 7	-1.39	0.0001	0.0318
100070632	AKR1E2	aldo-keto reductase family 1 member E2	protein-coding	83592	AKR1E2	aldo-keto reductase family 1 member E2	-1.93	0.0000	0.0179
100066458	ANKRD10	ankyrin repeat domain 10	protein-coding	55608	ANKRD10	ankyrin repeat domain 10	-0.66	0.0007	0.0998
100146761	ANO8	anoctamin 8	protein-coding	57719	ANO8	anoctamin 8	-1.24	0.0004	0.0763
100146445	APBA3	amyloid beta precursor protein binding family A member 3	protein-coding	9546	APBA3	amyloid beta precursor protein binding family A member 3	-0.62	0.0003	0.0708
100629225	ATP8B2	ATPase phospholipid transporting 8B2	protein-coding	57198	ATP8B2	ATPase phospholipid transporting 8B2	-0.59	0.0003	0.0676
100058097	C1QA	complement C1q A chain	protein-coding	712	C1QA	complement C1q A chain	-0.66	0.0006	0.0861
100071667	C1QB	complement C1q B chain	protein-coding	713	C1QB	complement C1q B chain	-0.74	0.0002	0.0562
100056159	CEP131	centrosomal protein 131	protein-coding	22994	CEP131	centrosomal protein 131	-0.85	0.0002	0.0529
100065271	CIAO3	cytosolic iron-sulfur assembly component 3	protein-coding	64428	CIAO3	cytosolic iron-sulfur assembly component 3	-0.57	0.0007	0.0998
100146651	CLCF1	cardiotrophin like cytokine factor 1	protein-coding	23529	CLCF1	cardiotrophin like cytokine factor 1	-2.76	0.0002	0.0529
100067832	CLK1	CDC like kinase 1	protein-coding	1195	CLK1	CDC like kinase 1	-0.94	0.0004	0.0796
100063546	CLK2	CDC like kinase 2	protein-coding	1196	CLK2	CDC like kinase 2	-0.69	0.0006	0.0861
100056083	COL16A1	collagen type XVI alpha 1 chain	protein-coding	1307	COL16A1	collagen type XVI alpha 1 chain	-1.72	0.0005	0.0847
100066148	COL4A1	collagen type IV alpha 1 chain	protein-coding	1282	COL4A1	collagen type IV alpha 1 chain	-1.16	0.0003	0.0658
100066264	COL4A2	collagen type IV alpha 2 chain	protein-coding	1284	COL4A2	collagen type IV alpha 2 chain	-1.17	0.0001	0.0354
100050035	COL6A1	collagen type VI alpha 1 chain	protein-coding	1291	COL6A1	collagen type VI alpha 1 chain	-1.84	0.0006	0.0897
100034138	CRISP2	cysteine-rich secretory protein 2	protein-coding	7180	CRISP2	cysteine rich secretory protein 2	-3.47	0.0000	0.0195
100069735	CYTH4	cytohesin 4	protein-coding	27128	CYTH4	cytohesin 4	-0.79	0.0004	0.0798
100065730	DENND1C	DENN domain containing 1C	protein-coding	79958	DENND1C	DENN domain containing 1C	-0.91	0.0006	0.0861
100061544	DLG4	discs large MAGUK scaffold protein 4	protein-coding	1742	DLG4	discs large MAGUK scaffold protein 4	-0.98	0.0003	0.0643
100057863	DNASE1L3	deoxyribonuclease 1 like 3	protein-coding	1776	DNASE1L3	deoxyribonuclease 1 like 3	-1.66	0.0001	0.0474
100057282	EHBP1L1	EH domain binding protein 1 like 1	protein-coding	254102	EHBP1L1	EH domain binding protein 1 like 1	-0.93	0.0001	0.0417

100057043	ENTPD6	ectonucleoside triphosphate diphosphohydrolase 6 (putative)	protein-coding	955	ENTPD6	ectonucleoside triphosphate diphosphohydrolase 6 (putative)	-0.70	0.0003	0.0620
102148534	ERFE	erythroferrone	protein-coding	151176	ERFE	erythroferrone	-2.36	0.0001	0.0471
100062182	FER1L5	fer-1 like family member 5	protein-coding	90342	FER1L5	fer-1 like family member 5	-1.77	0.0002	0.0591
100057614	FGFR1	fibroblast growth factor receptor 1	protein-coding	2260	FGFR1	fibroblast growth factor receptor 1	-1.03	0.0006	0.0883
100034189	FN1	fibronectin 1	protein-coding	2335	FN1	fibronectin 1	-2.21	0.0001	0.0472
100063926	GALNT16	polypeptide N-acetylgalactosaminyltransferase 16	protein-coding	57452	GALNT16	polypeptide N-acetylgalactosaminyltransferase 16	-2.24	0.0000	0.0065
100063638	GRAMD1B	GRAM domain containing 1B	protein-coding	57476	GRAMD1B	GRAM domain containing 1B	-0.61	0.0004	0.0763
100056429	HSD11B1	hydroxysteroid 11-beta dehydrogenase 1	protein-coding	3290	HSD11B1	hydroxysteroid 11-beta dehydrogenase 1	-2.26	0.0006	0.0861
100050779	HSPB8	heat shock protein family B (small) member 8	protein-coding	26353	HSPB8	heat shock protein family B (small) member 8	-2.76	0.0001	0.0417
100056218	IRF8	interferon regulatory factor 8	protein-coding	3394	IRF8	interferon regulatory factor 8	-0.79	0.0005	0.0861
100009709	ITGB3	integrin subunit beta 3	protein-coding	3690	ITGB3	integrin subunit beta 3	-2.58	0.0000	0.0130
100147451	JAK3	Janus kinase 3	protein-coding	3718	JAK3	Janus kinase 3	-0.66	0.0003	0.0673
100069672	KIF7	kinesin family member 7	protein-coding	374654	KIF7	kinesin family member 7	-0.82	0.0006	0.0867
100064430	LAT	linker for activation of T cells	protein-coding	27040	LAT	linker for activation of T cells	-0.93	0.0001	0.0417
100051856	LLGL1	LLGL1, scribble cell polarity complex component	protein-coding	3996	LLGL1	LLGL scribble cell polarity complex component 1	-0.50	0.0006	0.0861
100054029	LOC100054029	leukocyte immunoglobulin-like receptor subfamily A member 5	protein-coding				-1.26	0.0006	0.0883
100054448	LOC100054448	saoe class I histocompatibility antigen, A alpha chain	protein-coding	3105	HLA-A	major histocompatibility complex, class I, A	-1.26	0.0000	0.0024
100055483	LOC100055483	Ig mu chain C region membrane-bound form-like	other	3507	IGHM	immunoglobulin heavy constant mu	-1.82	0.0003	0.0609
100063097	LOC100063097	mitotic-spindle organizing protein 2B-like	protein-coding	80097	MZT2B	mitotic spindle organizing protein 2B	-1.70	0.0001	0.0423
100063831	LOC100063831	UL16-binding protein 1	protein-coding	79465	ULBP3	UL16 binding protein 3	-3.81	0.0000	0.0014
100070570	LOC100070570	aldo-keto reductase family 1 member C23	protein-coding	1109	AKR1C4	aldo-keto reductase family 1 member C4	-3.76	0.0002	0.0591
100073089	LOC100073089	ectonucleotide pyrophosphatase/phosphodiesterase family member 3	protein-coding	5169	ENPP3	ectonucleotide pyrophosphatase/phosphodiesterase 3	-1.26	0.0000	0.0200
100629324	LOC100629324	uncharacterized LOC100629324	protein-coding	55384	MEG3	maternally expressed 3	-2.85	0.0001	0.0417
102149846	LOC102149846	immunoglobulin heavy constant gamma 1-like	other	3500	IGHG1	immunoglobulin heavy constant gamma 1 (G1m marker)	-2.14	0.0005	0.0861
102150085	LOC102150085	immunoglobulin heavy constant gamma 1-like	other	3500	IGHG1	immunoglobulin heavy constant gamma 1 (G1m marker)	-2.67	0.0000	0.0240
102150790	LOC102150790	uncharacterized LOC102150790	ncRNA				-1.44	0.0000	0.0179
106781059	LOC106781059	uncharacterized LOC106781059	ncRNA				-1.46	0.0001	0.0327

106781303	LOC106781303	immunoglobulin heavy constant alpha 2-like	other	3493	IGHA1	immunoglobulin heavy constant alpha 1	-1.98	0.0005	0.0861
106781940	LOC106781940	uncharacterized LOC106781940	ncRNA				-1.45	0.0004	0.0762
106783330	LOC106783330	uncharacterized LOC106783330	protein-coding				-1.67	0.0001	0.0417
111767520	LOC111767520	uncharacterized LOC111767520	ncRNA				-1.60	0.0003	0.0669
111767890	LOC111767890	uncharacterized LOC111767890	ncRNA				-4.19	0.0000	0.0151
111768661	LOC111768661	translation initiation factor IF-2-like	protein-coding				-1.19	0.0006	0.0897
111768809	LOC111768809	uncharacterized LOC111768809	ncRNA				-2.47	0.0000	0.0113
111769010	LOC111769010	uncharacterized LOC111769010	ncRNA				-2.44	0.0000	0.0200
111771758	LOC111771758	GTPase IMAP family member 5-like	pseudo				-0.92	0.0001	0.0474
111774331	LOC111774331	uncharacterized LOC111774331	ncRNA				-1.17	0.0003	0.0620
111775969	LOC111775969	uncharacterized LOC111775969	ncRNA				-1.92	0.0005	0.0861
100067048	MCF2L	MCF.2 cell line derived transforming sequence like	protein-coding	23263	MCF2L	MCF.2 cell line derived transforming sequence like	-0.74	0.0001	0.0405
100066627	MICAL1	microtubule associated monooxygenase, calponin and LIM domain containing 1	protein-coding	64780	MICAL1	microtubule associated monooxygenase, calponin and LIM domain containing 1	-0.90	0.0000	0.0263
100068942	MMP25	matrix metallopeptidase 25	protein-coding	64386	MMP25	matrix metallopeptidase 25	-1.10	0.0003	0.0619
100057831	NAAA	N-acylethanolamine acid amidase	protein-coding	27163	NAAA	N-acylethanolamine acid amidase	-1.04	0.0003	0.0642
102148328	NMU	neuromedin U	protein-coding	10874	NMU	neuromedin U	-5.60	0.0002	0.0591
100062989	NPTX2	neuronal pentraxin 2	protein-coding	4885	NPTX2	neuronal pentraxin 2	-3.74	0.0001	0.0417
100056659	NUP210L	nucleoporin 210 like	protein-coding	91181	NUP210L	nucleoporin 210 like	-1.11	0.0002	0.0553
100050311	PDE10A	phosphodiesterase 10A	protein-coding	10846	PDE10A	phosphodiesterase 10A	-2.44	0.0000	0.0195
100052619	PDE2A	phosphodiesterase 2A	protein-coding	5138	PDE2A	phosphodiesterase 2A	-1.76	0.0004	0.0762
100057315	PLCB2	phospholipase C beta 2	protein-coding	5330	PLCB2	phospholipase C beta 2	-0.93	0.0005	0.0861
100058349	PLXNA3	plexin A3	protein-coding	55558	PLXNA3	plexin A3	-1.31	0.0005	0.0861
100146758	PNPLA2	patatin like phospholipase domain containing 2	protein-coding	57104	PNPLA2	patatin like phospholipase domain containing 2	-0.62	0.0006	0.0861
100071328	PREX1	phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 1	protein-coding	57580	PREX1	phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 1	-0.68	0.0006	0.0861
100629655	RAB44	RAB44, member RAS oncogene family	protein-coding	401258	RAB44	RAB44, member RAS oncogene family	-1.29	0.0005	0.0861
100034090	RYR1	ryanodine receptor 1	protein-coding	6261	RYR1	ryanodine receptor 1	-1.31	0.0004	0.0737
100053209	SDCBP2	syndecan binding protein 2	protein-coding	27111	SDCBP2	syndecan binding protein 2	-3.09	0.0000	0.0240
100072304	SLC37A2	solute carrier family 37 member 2	protein-coding	219855	SLC37A2	solute carrier family 37 member 2	-2.51	0.0002	0.0562
100067821	SLC39A8	solute carrier family 39 member 8	protein-coding	64116	SLC39A8	solute carrier family 39 member 8	-1.76	0.0007	0.0956
100056481	SLC8B1	solute carrier family 8 member B1	protein-coding	80024	SLC8B1	solute carrier family 8 member B1	-0.66	0.0003	0.0676
100054907	SNRNP70	small nuclear ribonucleoprotein U1 subunit 70	protein-coding	6625	SNRNP70	small nuclear ribonucleoprotein U1 subunit 70	-0.74	0.0002	0.0591
100050582	STAC	SH3 and cysteine rich domain	protein-coding	6769	STAC	SH3 and cysteine rich domain	-3.55	0.0000	0.0200

100052905	SYT12	synaptotagmin 12	protein-coding	91683	SYT12	synaptotagmin 12	-4.19	0.0001	0.0423
100050044	THBS2	thrombospondin 2	protein-coding	7058	THBS2	thrombospondin 2	-3.13	0.0000	0.0294
100050503	TIA1	TIA1 cytotoxic granule associated RNA binding protein	protein-coding	7072	TIA1	TIA1 cytotoxic granule associated RNA binding protein	-0.55	0.0006	0.0861
100146343	TNNT2	troponin T2, cardiac type	protein-coding	7139	TNNT2	troponin T2, cardiac type	-2.20	0.0001	0.0417
100059315	TNXB	tenascin XB	protein-coding	7148	TNXB	tenascin XB	-1.19	0.0000	0.0240
100051153	TOP3B	DNA topoisomerase III beta	protein-coding	8940	TOP3B	DNA topoisomerase III beta	-0.92	0.0002	0.0602
100069910	TPBG	trophoblast glycoprotein	protein-coding	7162	TPBG	trophoblast glycoprotein	-1.15	0.0004	0.0825
102150167	TPCN2	two pore segment channel 2	protein-coding	219931	TPCN2	two pore segment channel 2	-0.76	0.0002	0.0591
100069930	VGLL3	vestigial like family member 3	protein-coding	389136	VGLL3	vestigial like family member 3	-1.41	0.0006	0.0883
100066920	WDR90	WD repeat domain 90	protein-coding	197335	WDR90	WD repeat domain 90	-0.75	0.0005	0.0861
100055754	ZBP1	Z-DNA binding protein 1	protein-coding	81030	ZBP1	Z-DNA binding protein 1	-0.96	0.0000	0.0294
100064631	ZNF333	zinc finger protein 333	protein-coding	84449	ZNF333	zinc finger protein 333	-0.67	0.0007	0.0998
100060125	ZSWIM6	zinc finger SWIM-type containing 6	protein-coding	57688	ZSWIM6	zinc finger SWIM-type containing 6	-0.71	0.0006	0.0897
100054141	CRYL1	crystallin lambda 1	protein-coding	51084	CRYL1	crystallin lambda 1	0.89	0.0001	0.0338
100055161	CTSE	cathepsin E	protein-coding	1510	CTSE	cathepsin E	0.82	0.0002	0.0591
100063624	ELOVL2	ELOVL fatty acid elongase 2	protein-coding	54898	ELOVL2	ELOVL fatty acid elongase 2	0.96	0.0001	0.0474
100051778	LOC100051778	60S ribosomal protein L21	protein-coding	6144	RPL21	ribosomal protein L21	1.63	0.0000	0.0035
100052427	LOC100052427	60S ribosomal protein L26-like	pseudo				1.64	0.0000	0.0200
100065786	LOC100065786	40S ribosomal protein S17	protein-coding	6218	RPS17	ribosomal protein S17	1.46	0.0000	0.0014
100067178	LOC100067178	mesothelin	protein-coding	10232	MSLN	mesothelin	1.13	0.0005	0.0861
100072143	LOC100072143	centrin-4	protein-coding				0.80	0.0001	0.0318
100147232	LOC100147232	40S ribosomal protein S2 pseudogene	pseudo				1.80	0.0000	0.0012
111767704	LOC111767704	uncharacterized LOC111767704	ncRNA				0.91	0.0000	0.0179
100068904	MYCBP	MYC binding protein	protein-coding	26292	MYCBP	MYC binding protein	0.47	0.0002	0.0564
100629833	SDHAF4	succinate dehydrogenase complex assembly factor 4	protein-coding	135154	SDHAF4	succinate dehydrogenase complex assembly factor 4	0.45	0.0004	0.0763
100051264	SLC10A2	solute carrier family 10 member 2	protein-coding	6555	SLC10A2	solute carrier family 10 member 2	1.82	0.0002	0.0591
100060017	SLC16A5	solute carrier family 16 member 5	protein-coding	9121	SLC16A5	solute carrier family 16 member 5	0.57	0.0005	0.0861
100062703	SLC16A9	solute carrier family 16 member 9	protein-coding	220963	SLC16A9	solute carrier family 16 member 9	0.98	0.0001	0.0417
100630688	TOMM7	translocase of outer mitochondrial membrane 7	protein-coding	54543	TOMM7	translocase of outer mitochondrial membrane 7	0.52	0.0002	0.0591

Supplemental Table 2 Metascape Enrichment analysis- Genes lower in subfertile mares

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
1_Summary	KEGG Pathway	hsa04512	ECM-receptor interaction	-7.80	-3.34	7/82	1282,1284,1291,2335,3690,7058,7148,1307,712,713,43,64386,2260,3718,64116,11173,57104,57719,1742,7072,5330,55558,3996,57007,5138,6261,151176,1109,1196	COL4A1,COL4A2,COL6A1,FN1,ITGB3,THBS2,TNXB,COL16A1,C1QA,C1QB,ACHE,MMP25,FGFR1,JAK3,SLC39A8,ADAMTS7,PNPLA2,ANO8,DLG4,TIA1,PLCB2,PLXNA3,LLGL1,ACKR3,PDE2A,RYR1,ERFE,AKR1C4,CLK2
1_Member	KEGG Pathway	hsa04512	ECM-receptor interaction	-7.80	-3.34	7/82	1282,1284,1291,2335,3690,7058,7148	COL4A1,COL4A2,COL6A1,FN1,ITGB3,THBS2,TNXB
1_Member	Reactome Gene Sets	R-HSA-3000178	ECM proteoglycans	-6.55	-2.39	6/76	1282,1284,1291,2335,3690,7148	COL4A1,COL4A2,COL6A1,FN1,ITGB3,TNXB
1_Member	Reactome Gene Sets	R-HSA-216083	Integrin cell surface interactions	-6.26	-2.34	6/85	1282,1284,1291,1307,2335,3690	COL4A1,COL4A2,COL6A1,COL16A1,FN1,ITGB3
1_Member	GO Cellular Components	GO:0005581	collagen trimer	-6.20	-2.34	6/87	712,713,1282,1284,1291,1307	C1QA,C1QB,COL4A1,COL4A2,COL6A1,COL16A1
1_Member	GO Cellular Components	GO:0062023	collagen-containing extracellular matrix	-5.75	-2.04	10/410	43,712,713,1282,1284,1291,1307,2335,7058,7148	ACHE,C1QA,C1QB,COL4A1,COL4A2,COL6A1,COL16A1,FN1,THBS2,TNXB
1_Member	GO Molecular Functions	GO:0005201	extracellular matrix structural constituent	-5.72	-2.04	7/165	1282,1284,1291,1307,2335,7058,7148	COL4A1,COL4A2,COL6A1,COL16A1,FN1,THBS2,TNXB
1_Member	GO Cellular Components	GO:0031012	extracellular matrix	-5.56	-1.95	11/535	43,712,713,1282,1284,1291,1307,2335,7058,7148,64386	ACHE,C1QA,C1QB,COL4A1,COL4A2,COL6A1,COL16A1,FN1,THBS2,TNXB,MMP25
1_Member	KEGG Pathway	hsa04151	PI3K-Akt signaling pathway	-5.48	-1.93	9/342	1282,1284,1291,2260,2335,3690,3718,7058,7148	COL4A1,COL4A2,COL6A1,FGFR1,FN1,ITGB3,JAK3,THBS2,TNXB
1_Member	GO Biological Processes	GO:0030198	extracellular matrix organization	-5.18	-1.82	9/373	1282,1284,1291,1307,2335,3690,7148,64116,64386	COL4A1,COL4A2,COL6A1,COL16A1,FN1,ITGB3,TNXB,SLC39A8,MMP25
1_Member	KEGG Pathway	hsa04510	Focal adhesion	-5.18	-1.82	7/199	1282,1284,1291,2335,3690,7058,7148	COL4A1,COL4A2,COL6A1,FN1,ITGB3,THBS2,TNXB
1_Member	Canonical Pathways	M160	PID AVB3 INTEGRIN PATHWAY	-5.17	-1.82	5/75	1282,1291,1307,2335,3690	COL4A1,COL6A1,COL16A1,FN1,ITGB3
1_Member	Reactome Gene Sets	R-HSA-1474228	Degradation of the extracellular matrix	-4.99	-1.70	6/140	1282,1284,1291,1307,2335,64386	COL4A1,COL4A2,COL6A1,COL16A1,FN1,MMP25
1_Member	Reactome Gene Sets	R-HSA-1474244	Extracellular matrix organization	-4.95	-1.70	8/301	1282,1284,1291,1307,2335,3690,7148,64386	COL4A1,COL4A2,COL6A1,COL16A1,FN1,ITGB3,TNXB,MMP25
1_Member	GO Molecular Functions	GO:0030020	extracellular matrix structural constituent conferring tensile strength	-4.88	-1.68	4/41	1282,1284,1291,1307	COL4A1,COL4A2,COL6A1,COL16A1

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
1_Member	GO Cellular Components	GO:0005788	endoplasmic reticulum lumen	-4.88	-1.68	8/308	1282,1284,1291,1307,2335,11173,57104,57719	COL4A1,COL4A2,COL6A1,COL16A1,FN1,ADAMTS7,PNPLA2,ANO8
1_Member	Canonical Pathways	M3005	NABA COLLAGENS	-4.76	-1.66	4/44	1282,1284,1291,1307	COL4A1,COL4A2,COL6A1,COL16A1
1_Member	Reactome Gene Sets	R-HSA-8948216	Collagen chain trimerization	-4.76	-1.66	4/44	1282,1284,1291,1307	COL4A1,COL4A2,COL6A1,COL16A1
1_Member	GO Biological Processes	GO:0043062	extracellular structure organization	-4.69	-1.62	9/429	1282,1284,1291,1307,2335,3690,7148,64116,64386	COL4A1,COL4A2,COL6A1,COL16A1,FN1,ITGB3,TNXB,SLC39A8,MMP25
1_Member	GO Cellular Components	GO:0005604	basement membrane	-4.62	-1.56	5/97	43,1282,1284,2335,7058	ACHE,COL4A1,COL4A2,FN1,THBS2
1_Member	Reactome Gene Sets	R-HSA-186797	Signaling by PDGF	-4.28	-1.28	4/58	1282,1284,1291,7058	COL4A1,COL4A2,COL6A1,THBS2
1_Member	Canonical Pathways	M5884	NABA CORE MATRISOME	-4.27	-1.28	7/275	1282,1284,1291,1307,2335,7058,7148	COL4A1,COL4A2,COL6A1,COL16A1,FN1,THBS2,TNXB
1_Member	Reactome Gene Sets	R-HSA-3000171	Non-integrin membrane-ECM interactions	-4.25	-1.27	4/59	1282,1284,2335,3690	COL4A1,COL4A2,FN1,ITGB3
1_Member	Reactome Gene Sets	R-HSA-375165	NCAM signaling for neurite out-growth	-4.14	-1.19	4/63	1282,1284,1291,2260	COL4A1,COL4A2,COL6A1,FGFR1
1_Member	Reactome Gene Sets	R-HSA-9006934	Signaling by Receptor Tyrosine Kinases	-4.14	-1.19	9/506	1282,1284,1291,1742,2260,2335,3690,7058,7072	COL4A1,COL4A2,COL6A1,DLG4,FGFR1,FN1,ITGB3,THBS2,TIA1
1_Member	Hallmark Gene Sets	M5930	HALLMARK EPITHELIAL MESENCHYMAL TRANSITION	-4.12	-1.19	6/200	1282,1284,1307,2335,3690,7058	COL4A1,COL4A2,COL16A1,FN1,ITGB3,THBS2
1_Member	Reactome Gene Sets	R-HSA-1442490	Collagen degradation	-4.11	-1.19	4/64	1282,1284,1291,1307	COL4A1,COL4A2,COL6A1,COL16A1
1_Member	GO Biological Processes	GO:0001704	formation of primary germ layer	-4.08	-1.18	5/126	1284,1291,2260,2335,3690	COL4A2,COL6A1,FGFR1,FN1,ITGB3
1_Member	Canonical Pathways	M18	PID INTEGRIN1 PATHWAY	-4.06	-1.17	4/66	1282,1291,2335,7058	COL4A1,COL6A1,FN1,THBS2
1_Member	Reactome Gene Sets	R-HSA-1650814	Collagen biosynthesis and modifying enzymes	-4.03	-1.16	4/67	1282,1284,1291,1307	COL4A1,COL4A2,COL6A1,COL16A1
1_Member	Reactome Gene Sets	R-HSA-1474290	Collagen formation	-3.54	-0.81	4/90	1282,1284,1291,1307	COL4A1,COL4A2,COL6A1,COL16A1
1_Member	KEGG Pathway	hsa05146	Amoebiasis	-3.43	-0.78	4/96	1282,1284,2335,5330	COL4A1,COL4A2,FN1,PLCB2
1_Member	Canonical Pathways	M5887	NABA BASEMENT MEMBRANES	-3.42	-0.78	3/40	1282,1284,1291	COL4A1,COL4A2,COL6A1
1_Member	KEGG Pathway	hsa04933	AGE-RAGE signaling pathway in diabetic complications	-3.38	-0.76	4/99	1282,1284,2335,5330	COL4A1,COL4A2,FN1,PLCB2

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
1_Member	Reactome Gene Sets	R-HSA-419037	NCAM1 interactions	-3.36	-0.74	3/42	1282,1284,1291	COL4A1,COL4A2,COL6A1
1_Member	Canonical Pathways	M53	PID INTEGRIN3 PATHWAY	-3.33	-0.72	3/43	1282,2335,3690	COL4A1,FN1,ITGB3
1_Member	GO Biological Processes	GO:0035987	endodermal cell differentiation	-3.27	-0.68	3/45	1284,1291,2335	COL4A2,COL6A1,FN1
1_Member	GO Biological Processes	GO:0007369	gastrulation	-3.27	-0.68	5/188	1284,1291,2260,2335,3690	COL4A2,COL6A1,FGFR1,FN1,ITGB3
1_Member	Canonical Pathways	M198	PID SYNDECAN 1 PATHWAY	-3.24	-0.67	3/46	1282,1291,1307	COL4A1,COL6A1,COL16A1
1_Member	GO Biological Processes	GO:0001706	endoderm formation	-3.04	-0.52	3/54	1284,1291,2335	COL4A2,COL6A1,FN1
1_Member	Reactome Gene Sets	R-HSA-2022090	Assembly of collagen fibrils and other multimeric structures	-2.88	-0.43	3/61	1282,1284,1291	COL4A1,COL4A2,COL6A1
1_Member	GO Molecular Functions	GO:0019838	growth factor binding	-2.81	-0.39	4/141	1282,1291,2260,3690	COL4A1,COL6A1,FGFR1,ITGB3
1_Member	GO Biological Processes	GO:0071230	cellular response to amino acid stimulus	-2.73	-0.35	3/69	1282,1291,1307	COL4A1,COL6A1,COL16A1
1_Member	GO Molecular Functions	GO:0005518	collagen binding	-2.71	-0.33	3/70	43,1291,2335	ACHE,COL6A1,FN1
1_Member	GO Biological Processes	GO:0007492	endoderm development	-2.58	-0.22	3/78	1284,1291,2335	COL4A2,COL6A1,FN1
1_Member	KEGG Pathway	hsa05222	Small cell lung cancer	-2.49	-0.15	3/84	1282,1284,2335	COL4A1,COL4A2,FN1
1_Member	Reactome Gene Sets	R-HSA-422475	Axon guidance	-2.48	-0.15	7/552	1282,1284,1291,1742,2260,3690,55558	COL4A1,COL4A2,COL6A1,DLG4,FGFR1,ITGB3,PLXNA3
1_Member	GO Molecular Functions	GO:0005198	structural molecule activity	-2.44	-0.12	8/714	1282,1284,1291,1307,2335,3996,7058,7148	COL4A1,COL4A2,COL6A1,COL16A1,FN1,LLGL1,THBS2,TNXB
1_Member	KEGG Pathway	hsa04974	Protein digestion and absorption	-2.40	-0.09	3/90	1282,1284,1291	COL4A1,COL4A2,COL6A1
1_Member	Reactome Gene Sets	R-HSA-9675108	Nervous system development	-2.37	-0.07	7/577	1282,1284,1291,1742,2260,3690,55558	COL4A1,COL4A2,COL6A1,DLG4,FGFR1,ITGB3,PLXNA3
1_Member	GO Biological Processes	GO:0001525	angiogenesis	-2.30	-0.01	7/594	1282,1284,2260,2335,3690,7058,57007	COL4A1,COL4A2,FGFR1,FN1,ITGB3,THBS2,ACKR3
1_Member	GO Biological Processes	GO:0071417	cellular response to organonitrogen compound	-2.29	-0.01	7/597	1282,1291,1307,3718,5138,6261,151176	COL4A1,COL6A1,COL16A1,JAK3,PDE2A,RYR1,ERFE
1_Member	GO Biological Processes	GO:0001568	blood vessel development	-2.22	0.00	8/777	1282,1284,2260,2335,3690,5138,7058,57007	COL4A1,COL4A2,FGFR1,FN1,ITGB3,PDE2A,THBS2,ACKR3
1_Member	GO Biological Processes	GO:0071229	cellular response to acid chemical	-2.17	0.00	4/212	1109,1282,1291,1307	AKR1C4,COL4A1,COL6A1,COL16A1

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
1_Member	GO Biological Processes	GO:0043200	response to amino acid	-2.12	0.00	3/114	1282,1291,1307	COL4A1,COL6A1,COL16A1
1_Member	GO Biological Processes	GO:0001101	response to acid chemical	-2.12	0.00	5/347	1109,1196,1282,1291,1307	AKR1C4,CLK2,COL4A1,COL6A1,COL16A1
1_Member	GO Biological Processes	GO:1901699	cellular response to nitrogen compound	-2.06	0.00	7/660	1282,1291,1307,3718,5138,6261,151176	COL4A1,COL6A1,COL16A1,JAK3,PDE2A,RYR1,ERFE
2_Summary	GO Biological Processes	GO:0002449	lymphocyte mediated immunity	-5.26	-1.82	9/364	712,713,3105,3493,3500,3507,23529,28432,79465,3718,1742,5169,6769,57007,27040,2335,57580,3394,3690	C1QA,C1QB,HLA-A,IGHA1,IGHG1,IGHM,CLCF1,IGHV3-35,ULBP3,JAK3,DLG4,ENPP3,STAC,ACKR3,LAT,FN1,PREX1,IRF8,ITGB3
2_Member	GO Biological Processes	GO:0002449	lymphocyte mediated immunity	-5.26	-1.82	9/364	712,713,3105,3493,3500,3507,23529,28432,79465	C1QA,C1QB,HLA-A,IGHA1,IGHG1,IGHM,CLCF1,IGHV3-35,ULBP3
2_Member	GO Biological Processes	GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobuli	-5.24	-1.82	9/366	712,713,3105,3493,3500,3507,3718,23529,28432	C1QA,C1QB,HLA-A,IGHA1,IGHG1,IGHM,JAK3,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0006958	complement activation, classical pathway	-4.97	-1.70	6/141	712,713,3493,3500,3507,28432	C1QA,C1QB,IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0016064	immunoglobulin mediated immune response	-4.83	-1.66	7/225	712,713,3493,3500,3507,23529,28432	C1QA,C1QB,IGHA1,IGHG1,IGHM,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0019724	B cell mediated immunity	-4.79	-1.66	7/228	712,713,3493,3500,3507,23529,28432	C1QA,C1QB,IGHA1,IGHG1,IGHM,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0002455	humoral immune response mediated by circulating immunoglobulin	-4.75	-1.66	6/154	712,713,3493,3500,3507,28432	C1QA,C1QB,IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0006956	complement activation	-4.40	-1.36	6/178	712,713,3493,3500,3507,28432	C1QA,C1QB,IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Cellular Components	GO:0098552	side of membrane	-4.32	-1.29	10/599	1742,3105,3493,3500,3507,5169,6769,28432,57007,79465	DLG4,HLA-A,IGHA1,IGHG1,IGHM,ENPP3,STAC,IGHV3-35,ACKR3,ULBP3
2_Member	GO Biological Processes	GO:0072376	protein activation cascade	-4.11	-1.19	6/201	712,713,3493,3500,3507,28432	C1QA,C1QB,IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Cellular Components	GO:0042571	immunoglobulin complex, circulating	-3.82	-0.95	4/76	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0002250	adaptive immune response	-3.78	-0.93	10/696	712,713,3105,3493,3500,3507,3718,23529,27040,28432	C1QA,C1QB,HLA-A,IGHA1,IGHG1,IGHM,JAK3,CLCF1,LAT,IGHV3-35
2_Member	GO Cellular Components	GO:0072562	blood microparticle	-3.76	-0.93	5/147	713,2335,3493,3500,3507	C1QB,FN1,IGHA1,IGHG1,IGHM

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
2_Member	GO Molecular Functions	GO:0034987	immunoglobulin receptor binding	-3.73	-0.92	4/80	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0050871	positive regulation of B cell activation	-3.72	-0.92	5/150	3493,3500,3507,23529,28432	IGHA1,IGHG1,IGHM,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0051251	positive regulation of lymphocyte activation	-3.66	-0.86	7/346	3105,3493,3500,3507,3718,23529,28432	HLA-A,IGHA1,IGHG1,IGHM,JAK3,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0002694	regulation of leukocyte activation	-3.65	-0.86	9/588	3105,3493,3500,3507,3718,5169,23529,27040,28432	HLA-A,IGHA1,IGHG1,IGHM,JAK3,ENPP3,CLCF1,LAT,IGHV3-35
2_Member	GO Biological Processes	GO:0006910	phagocytosis, recognition	-3.58	-0.83	4/88	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0046649	lymphocyte activation	-3.53	-0.81	10/748	3105,3493,3500,3507,3718,23529,27040,28432,57580,79465	HLA-A,IGHA1,IGHG1,IGHM,JAK3,CLCF1,LAT,IGHV3-35,PREX1,ULBP3
2_Member	GO Molecular Functions	GO:0003823	antigen binding	-3.48	-0.78	5/169	3105,3493,3500,3507,28432	HLA-A,IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0050865	regulation of cell activation	-3.42	-0.78	9/632	3105,3493,3500,3507,3718,5169,23529,27040,28432	HLA-A,IGHA1,IGHG1,IGHM,JAK3,ENPP3,CLCF1,LAT,IGHV3-35
2_Member	GO Biological Processes	GO:0051249	regulation of lymphocyte activation	-3.41	-0.78	8/501	3105,3493,3500,3507,3718,23529,27040,28432	HLA-A,IGHA1,IGHG1,IGHM,JAK3,CLCF1,LAT,IGHV3-35
2_Member	GO Biological Processes	GO:0002696	positive regulation of leukocyte activation	-3.34	-0.72	7/391	3105,3493,3500,3507,3718,23529,28432	HLA-A,IGHA1,IGHG1,IGHM,JAK3,CLCF1,IGHV3-35
2_Member	GO Cellular Components	GO:0009897	external side of plasma membrane	-3.27	-0.68	7/402	3493,3500,3507,5169,28432,57007,79465	IGHA1,IGHG1,IGHM,ENPP3,IGHV3-35,ACKR3,ULBP3
2_Member	GO Biological Processes	GO:0050867	positive regulation of cell activation	-3.24	-0.67	7/406	3105,3493,3500,3507,3718,23529,28432	HLA-A,IGHA1,IGHG1,IGHM,JAK3,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0050864	regulation of B cell activation	-3.21	-0.66	5/194	3493,3500,3507,23529,28432	IGHA1,IGHG1,IGHM,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0042113	B cell activation	-3.03	-0.52	6/320	3493,3500,3507,3718,23529,28432	IGHA1,IGHG1,IGHM,JAK3,CLCF1,IGHV3-35
2_Member	GO Biological Processes	GO:0050851	antigen receptor-mediated signaling pathway	-3.03	-0.52	6/320	3105,3493,3500,3507,27040,28432	HLA-A,IGHA1,IGHG1,IGHM,LAT,IGHV3-35
2_Member	GO Biological Processes	GO:0006911	phagocytosis, engulfment	-3.03	-0.52	4/123	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0050853	B cell receptor signaling pathway	-2.91	-0.44	4/132	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0099024	plasma membrane invagination	-2.91	-0.44	4/132	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
2_Member	GO Biological Processes	GO:0010324	membrane invagination	-2.82	-0.39	4/140	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0006959	humoral immune response	-2.76	-0.36	6/361	712,713,3493,3500,3507,28432	C1QA,C1QB,IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0006909	phagocytosis	-2.70	-0.32	6/372	3394,3493,3500,3507,3690,28432	IRF8,IGHA1,IGHG1,IGHM,ITGB3,IGHV3-35
2_Member	GO Cellular Components	GO:0019814	immunoglobulin complex	-2.58	-0.22	4/163	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0002253	activation of immune response	-2.38	-0.07	8/731	712,713,3105,3493,3500,3507,27040,28432	C1QA,C1QB,HLA-A,IGHA1,IGHG1,IGHM,LAT,IGHV3-35
2_Member	GO Biological Processes	GO:0002429	immune response-activating cell surface receptor signaling pathway	-2.17	0.00	6/476	3105,3493,3500,3507,27040,28432	HLA-A,IGHA1,IGHG1,IGHM,LAT,IGHV3-35
2_Member	GO Biological Processes	GO:0042742	defense response to bacterium	-2.13	0.00	5/344	3394,3493,3500,3507,28432	IRF8,IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0008037	cell recognition	-2.13	0.00	4/218	3493,3500,3507,28432	IGHA1,IGHG1,IGHM,IGHV3-35
2_Member	GO Biological Processes	GO:0002768	immune response-regulating cell surface receptor signaling pathway	-2.04	0.00	6/509	3105,3493,3500,3507,27040,28432	HLA-A,IGHA1,IGHG1,IGHM,LAT,IGHV3-35
3_Summary	GO Biological Processes	GO:0007204	positive regulation of cytosolic calcium ion concentration	-3.77	-0.93	7/332	1742,5330,6261,10874,57007,80024,219931,64116,151176,1291,6769,3690,57104,2260,91683,57198,57719,7139	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC8B1,TPCN2,SLC39A8,ERFE,COL6A1,STAC,ITGB3,PNPLA2,FGFR1,SYT12,ATP8B2,ANO8,TNNT2
3_Member	GO Biological Processes	GO:0007204	positive regulation of cytosolic calcium ion concentration	-3.77	-0.93	7/332	1742,5330,6261,10874,57007,80024,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0006875	cellular metal ion homeostasis	-3.61	-0.85	9/596	1742,5330,6261,10874,57007,64116,80024,151176,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,ERFE,TPCN2
3_Member	GO Biological Processes	GO:0051480	regulation of cytosolic calcium ion concentration	-3.46	-0.78	7/373	1742,5330,6261,10874,57007,80024,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0072503	cellular divalent inorganic cation homeostasis	-3.38	-0.76	8/507	1742,5330,6261,10874,57007,64116,80024,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0072507	divalent inorganic cation homeostasis	-3.27	-0.68	8/526	1742,5330,6261,10874,57007,64116,80024,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0055065	metal ion homeostasis	-3.25	-0.67	9/668	1742,5330,6261,10874,57007,64116,80024,151176,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,ERFE,TPCN2
3_Member	GO Biological Processes	GO:0030003	cellular cation homeostasis	-3.23	-0.67	9/672	1742,5330,6261,10874,57007,64116,80024,151176,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,ERFE,TPCN2

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
3_Member	GO Biological Processes	GO:0006873	cellular ion homeostasis	-3.17	-0.62	9/685	1742,5330,6261,10874,57007,64116,80024,151176,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,ERFE,TPCN2
3_Member	GO Biological Processes	GO:0051209	release of sequestered calcium ion into cytosol	-2.92	-0.44	4/131	5330,6261,80024,219931	PLCB2,RYR1,SLC8B1,TPCN2
3_Member	GO Cellular Components	GO:0042383	sarcolemma	-2.91	-0.44	4/132	1291,6261,6769,80024	COL6A1,RYR1,STAC,SLC8B1
3_Member	GO Biological Processes	GO:0051283	negative regulation of sequestering of calcium ion	-2.91	-0.44	4/132	5330,6261,80024,219931	PLCB2,RYR1,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0055080	cation homeostasis	-2.90	-0.43	9/748	1742,5330,6261,10874,57007,64116,80024,151176,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,ERFE,TPCN2
3_Member	GO Biological Processes	GO:0051282	regulation of sequestering of calcium ion	-2.89	-0.43	4/134	5330,6261,80024,219931	PLCB2,RYR1,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0051235	maintenance of location	-2.87	-0.42	6/343	3690,5330,6261,57104,80024,219931	ITGB3,PLCB2,RYR1,PNPLA2,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0006874	cellular calcium ion homeostasis	-2.86	-0.42	7/472	1742,5330,6261,10874,57007,80024,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0098771	inorganic ion homeostasis	-2.86	-0.42	9/759	1742,5330,6261,10874,57007,64116,80024,151176,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC39A8,SLC8B1,ERFE,TPCN2
3_Member	GO Biological Processes	GO:0051208	sequestering of calcium ion	-2.85	-0.42	4/137	5330,6261,80024,219931	PLCB2,RYR1,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0055074	calcium ion homeostasis	-2.80	-0.39	7/484	1742,5330,6261,10874,57007,80024,219931	DLG4,PLCB2,RYR1,NMU,ACKR3,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0097553	calcium ion transmembrane import into cytosol	-2.74	-0.35	4/147	5330,6261,80024,219931	PLCB2,RYR1,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0060402	calcium ion transport into cytosol	-2.55	-0.20	4/166	5330,6261,80024,219931	PLCB2,RYR1,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0043269	regulation of ion transport	-2.49	-0.15	8/700	1742,2260,3690,6769,80024,91683,151176,219931	DLG4,FGFR1,ITGB3,STAC,SLC8B1,SYT12,ERFE,TPCN2
3_Member	GO Biological Processes	GO:0060401	cytosolic calcium ion transport	-2.42	-0.11	4/180	5330,6261,80024,219931	PLCB2,RYR1,SLC8B1,TPCN2
3_Member	Reactome Gene Sets	R-HSA-983712	Ion channel transport	-2.37	-0.07	4/186	6261,57198,57719,219931	RYR1,ATP8B2,ANO8,TPCN2
3_Member	GO Biological Processes	GO:0070588	calcium ion transmembrane transport	-2.21	0.00	5/329	5330,6261,6769,80024,219931	PLCB2,RYR1,STAC,SLC8B1,TPCN2
3_Member	Reactome Gene Sets	R-HSA-2672351	Stimuli-sensing channels	-2.17	0.00	3/109	6261,57719,219931	RYR1,ANO8,TPCN2
3_Member	GO Biological Processes	GO:0070838	divalent metal ion transport	-2.09	0.00	6/495	5330,6261,6769,64116,80024,219931	PLCB2,RYR1,STAC,SLC39A8,SLC8B1,TPCN2

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
3_Member	GO Biological Processes	GO:0072511	divalent inorganic cation transport	-2.07	0.00	6/501	5330,6261,6769,64116,80024,219931	PLCB2,RYR1,STAC,SLC39A8,SLC8B1,TPCN2
3_Member	GO Biological Processes	GO:0006936	muscle contraction	-2.06	0.00	5/359	6261,6769,7139,10874,219931	RYR1,STAC,TNNT2,NMU,TPCN2
4_Summary	GO Molecular Functions	GO:0005539	glycosaminoglycan binding	-3.76	-0.93	6/233	1307,2260,2335,3507,7058,7148,3690,57580,3493,80024	regu
4_Member	GO Molecular Functions	GO:0005539	glycosaminoglycan binding	-3.76	-0.93	6/233	1307,2260,2335,3507,7058,7148	COL16A1,FGFR1,FN1,IGHM,THBS2,TNXB
4_Member	GO Molecular Functions	GO:0008201	heparin binding	-3.48	-0.78	5/169	1307,2260,2335,7058,7148	COL16A1,FGFR1,FN1,THBS2,TNXB
4_Member	GO Molecular Functions	GO:0005178	integrin binding	-2.85	-0.42	4/137	1307,2335,3690,7148	COL16A1,FN1,ITGB3,TNXB
4_Member	GO Molecular Functions	GO:1901681	sulfur compound binding	-2.68	-0.31	5/255	1307,2260,2335,7058,7148	COL16A1,FGFR1,FN1,THBS2,TNXB
4_Member	GO Cellular Components	GO:0031091	platelet alpha granule	-2.39	-0.08	3/91	2335,3690,7058	FN1,ITGB3,THBS2
4_Member	GO Biological Processes	GO:0034446	substrate adhesion-dependent cell spreading	-2.30	-0.01	3/98	2335,3690,57580	FN1,ITGB3,PREX1
4_Member	GO Biological Processes	GO:0050900	leukocyte migration	-2.06	0.00	6/504	2335,3493,3507,3690,57580,80024	FN1,IGHA1,IGHM,ITGB3,PREX1,SLC8B1
4_Member	GO Biological Processes	GO:0010811	positive regulation of cell-substrate adhesion	-2.06	0.00	3/120	1307,2335,57580	COL16A1,FN1,PREX1
5_Summary	GO Biological Processes	GO:0002697	regulation of immune effector process	-3.62	-0.85	8/466	712,713,1776,3105,3500,3718,5169,23529,79465	C1QA,C1QB,DNASE1L3,HLA-A,IGHG1,JAK3,ENPP3,CLCF1,ULBP3
5_Member	GO Biological Processes	GO:0002697	regulation of immune effector process	-3.62	-0.85	8/466	712,713,1776,3105,3500,3718,5169,23529	C1QA,C1QB,DNASE1L3,HLA-A,IGHG1,JAK3,ENPP3,CLCF1
5_Member	GO Biological Processes	GO:0002703	regulation of leukocyte mediated immunity	-2.22	0.00	4/206	1776,3105,3718,23529	DNASE1L3,HLA-A,JAK3,CLCF1
5_Member	GO Biological Processes	GO:0001909	leukocyte mediated cytotoxicity	-2.16	0.00	3/110	1776,3105,79465	DNASE1L3,HLA-A,ULBP3
6_Summary	Reactome Gene Sets	R-HSA-5673001	RAF/MAP kinase cascade	-3.61	-0.85	6/248	1742,2260,2335,3690,3718,27040,1195,1196,23529,3500,1307,5330	DLG4,FGFR1,FN1,ITGB3,JAK3,LAT,CLK1,CLK2,CLCF1,IGHG1,COL16A1,PLCB2
6_Member	Reactome Gene Sets	R-HSA-5673001	RAF/MAP kinase cascade	-3.61	-0.85	6/248	1742,2260,2335,3690,3718,27040	DLG4,FGFR1,FN1,ITGB3,JAK3,LAT
6_Member	Reactome Gene Sets	R-HSA-5684996	MAPK1/MAPK3 signaling	-3.56	-0.81	6/254	1742,2260,2335,3690,3718,27040	DLG4,FGFR1,FN1,ITGB3,JAK3,LAT
6_Member	GO Biological Processes	GO:0018108	peptidyl-tyrosine phosphorylation	-3.47	-0.78	7/371	1195,1196,1742,2260,3690,3718,23529	CLK1,CLK2,DLG4,FGFR1,ITGB3,JAK3,CLCF1

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
6_Member	GO Biological Processes	GO:0018212	peptidyl-tyrosine modification	-3.45	-0.78	7/374	1195,1196,1742,2260,3690,3718,23529	CLK1,CLK2,DLG4,FGFR1,ITGB3,JAK3,CLCF1
6_Member	Reactome Gene Sets	R-HSA-9607240	FLT3 Signaling	-3.45	-0.78	6/266	1742,2260,2335,3690,3718,27040	DLG4,FGFR1,FN1,ITGB3,JAK3,LAT
6_Member	Reactome Gene Sets	R-HSA-5683057	MAPK family signaling cascades	-3.23	-0.67	6/293	1742,2260,2335,3690,3718,27040	DLG4,FGFR1,FN1,ITGB3,JAK3,LAT
6_Member	GO Molecular Functions	GO:0004713	protein tyrosine kinase activity	-2.83	-0.40	4/139	1195,1196,2260,3718	CLK1,CLK2,FGFR1,JAK3
6_Member	Canonical Pathways	M28	PID IL4 2PATHWAY	-2.80	-0.39	3/65	3500,3690,3718	IGHG1,ITGB3,JAK3
6_Member	GO Biological Processes	GO:0007229	integrin-mediated signaling pathway	-2.19	0.00	3/107	1307,3690,27040	COL16A1,ITGB3,LAT
6_Member	KEGG Pathway	hsa04015	Rap1 signaling pathway	-2.19	0.00	4/210	2260,3690,5330,27040	FGFR1,ITGB3,PLCB2,LAT
6_Member	Reactome Gene Sets	R-HSA-373760	L1CAM interactions	-2.07	0.00	3/119	1742,2260,3690	DLG4,FGFR1,ITGB3
7_Summary	GO Biological Processes	GO:0015698	inorganic anion transport	-3.47	-0.78	5/170	2260,5169,57719,64116,219855,1109,57198,151176	FGFR1,ENPP3,ANO8,SLC39A8,SLC37A2,AKR1C4,ATP8B2,ERFE
7_Member	GO Biological Processes	GO:0015698	inorganic anion transport	-3.47	-0.78	5/170	2260,5169,57719,64116,219855	FGFR1,ENPP3,ANO8,SLC39A8,SLC37A2
7_Member	GO Biological Processes	GO:0006820	anion transport	-2.74	-0.35	8/637	1109,2260,5169,57198,57719,64116,151176,219855	AKR1C4,FGFR1,ENPP3,ATP8B2,ANO8,SLC39A8,ERFE,SLC37A2
8_Summary	GO Molecular Functions	GO:0008081	phosphoric diester hydrolase activity	-3.45	-0.78	4/95	5138,5169,5330,10846,955	PDE2A,ENPP3,PLCB2,PDE10A,ENTPD6
8_Member	GO Molecular Functions	GO:0008081	phosphoric diester hydrolase activity	-3.45	-0.78	4/95	5138,5169,5330,10846	PDE2A,ENPP3,PLCB2,PDE10A
8_Member	KEGG Pathway	hsa00230	Purine metabolism	-2.48	-0.15	4/174	955,5138,5169,10846	ENTPD6,PDE2A,ENPP3,PDE10A
9_Summary	GO Biological Processes	GO:0031623	receptor internalization	-3.15	-0.62	4/114	43,1742,3690,57007,3493	ACHE,DLG4,ITGB3,ACKR3,IGHA1
9_Member	GO Biological Processes	GO:0031623	receptor internalization	-3.15	-0.62	4/114	43,1742,3690,57007	ACHE,DLG4,ITGB3,ACKR3
9_Member	GO Biological Processes	GO:0043112	receptor metabolic process	-2.29	-0.01	4/196	43,1742,3690,57007	ACHE,DLG4,ITGB3,ACKR3
9_Member	GO Biological Processes	GO:0006898	receptor-mediated endocytosis	-2.22	0.00	5/327	43,1742,3493,3690,57007	ACHE,DLG4,IGHA1,ITGB3,ACKR3
10_Summary	Hallmark Gene Sets	M5909	HALLMARK MYOGENESIS	-3.15	-0.62	5/200	43,1284,6261,7139,26353,955,91683,712,1742,1776,5169,5330,7058,401258	ACHE,COL4A2,RYR1,TNNT2,HSPB8,ENTPD6,SYT12,C1QA,DLG4,DNASE1L3,ENPP3,PLCB2,THBS2,RAB44

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
10_Member	Hallmark Gene Sets	M5909	HALLMARK MYOGENESIS	-3.15	-0.62	5/200	43,1284,6261,7139,26353	ACHE,COL4A2,RYR1,TNNT2,HSPB8
10_Member	GO Biological Processes	GO:0051592	response to calcium ion	-2.75	-0.35	4/146	955,6261,7139,91683	ENTPD6,RYR1,TNNT2,SYT12
10_Member	GO Biological Processes	GO:0010038	response to metal ion	-2.75	-0.35	6/363	712,955,1742,6261,7139,91683	C1QA,ENTPD6,DLG4,RYR1,TNNT2,SYT12
10_Member	GO Molecular Functions	GO:0005509	calcium ion binding	-2.47	-0.15	8/705	1776,5169,5330,6261,7058,7139,91683,401258	DNASE1L3,ENPP3,PLCB2,RYR1,THBS2,TNNT2,SYT12,RAB44
11_Summary	GO Biological Processes	GO:0050808	synapse organization	-3.09	-0.56	7/431	43,712,713,1282,1742,7058,7162,2260,2335,23529,55558	ACHE,C1QA,C1QB,COL4A1,DLG4,THBS2,TPBG,FGFR1,FN1,CLCF1,PLXNA3
11_Member	GO Biological Processes	GO:0050808	synapse organization	-3.09	-0.56	7/431	43,712,713,1282,1742,7058,7162	ACHE,C1QA,C1QB,COL4A1,DLG4,THBS2,TPBG
11_Member	GO Biological Processes	GO:0051962	positive regulation of nervous system development	-2.48	-0.15	7/552	1742,2260,2335,7058,7162,23529,55558	DLG4,FGFR1,FN1,THBS2,TPBG,CLCF1,PLXNA3
12_Summary	GO Molecular Functions	GO:0005543	phospholipid binding	-3.07	-0.55	7/434	3507,23263,27111,57476,57580,79958,91683,3290,27128	IGHM,MCF2L,SDCBP2,GRAMD1B,PREX1,DENND1C,SYT12,HSD11B1,CYTH4
12_Member	GO Molecular Functions	GO:0005543	phospholipid binding	-3.07	-0.55	7/434	3507,23263,27111,57476,57580,79958,91683	IGHM,MCF2L,SDCBP2,GRAMD1B,PREX1,DENND1C,SYT12
12_Member	GO Molecular Functions	GO:0008289	lipid binding	-2.87	-0.42	9/755	3290,3507,23263,27111,27128,57476,57580,79958,91683	HSD11B1,IGHM,MCF2L,SDCBP2,CYTH4,GRAMD1B,PREX1,DENND1C,SYT12
12_Member	GO Molecular Functions	GO:0005085	guanyl-nucleotide exchange factor activity	-2.14	0.00	4/217	23263,27128,57580,79958	MCF2L,CYTH4,PREX1,DENND1C
13_Summary	Reactome Gene Sets	R-HSA-6794361	Neurexins and neuroligins	-3.01	-0.52	3/55	1742,9546,91683,43,712,5330,27163	DLG4,APBA3,SYT12,ACHE,C1QA,PLCB2,NAAA
13_Member	Reactome Gene Sets	R-HSA-6794361	Neurexins and neuroligins	-3.01	-0.52	3/55	1742,9546,91683	DLG4,APBA3,SYT12
13_Member	GO Molecular Functions	GO:0001540	amyloid-beta binding	-2.53	-0.18	3/81	43,712,9546	ACHE,C1QA,APBA3
13_Member	Reactome Gene Sets	R-HSA-112316	Neuronal System	-2.49	-0.15	6/410	43,1742,5330,9546,27163,91683	ACHE,DLG4,PLCB2,APBA3,NAAA,SYT12
13_Member	Reactome Gene Sets	R-HSA-6794362	Protein-protein interactions at synapses	-2.46	-0.14	3/86	1742,9546,91683	DLG4,APBA3,SYT12
14_Summary	GO Biological Processes	GO:0002526	acute inflammatory response	-2.95	-0.45	5/222	712,713,1776,2335,3500,3394,5330,5138,5169	C1QA,C1QB,DNASE1L3,FN1,IGHG1,IRF8,PLCB2,PDE2A,ENPP3
14_Member	GO Biological Processes	GO:0002526	acute inflammatory response	-2.95	-0.45	5/222	712,713,1776,2335,3500	C1QA,C1QB,DNASE1L3,FN1,IGHG1

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
14_Member	KEGG Pathway	hsa05133	Pertussis	-2.61	-0.24	3/76	712,713,3394	C1QA,C1QB,IRF8
14_Member	GO Biological Processes	GO:0002673	regulation of acute inflammatory response	-2.60	-0.23	4/161	712,713,1776,3500	C1QA,C1QB,DNASE1L3,IGHG1
14_Member	KEGG Pathway	hsa05142	Chagas disease (American trypanosomiasis)	-2.25	0.00	3/102	712,713,5330	C1QA,C1QB,PLCB2
14_Member	GO Biological Processes	GO:0030449	regulation of complement activation	-2.10	0.00	3/116	712,713,3500	C1QA,C1QB,IGHG1
14_Member	GO Biological Processes	GO:2000257	regulation of protein activation cascade	-2.09	0.00	3/117	712,713,3500	C1QA,C1QB,IGHG1
14_Member	GO Biological Processes	GO:0050727	regulation of inflammatory response	-2.04	0.00	6/508	712,713,1776,3500,5138,5169	C1QA,C1QB,DNASE1L3,IGHG1,PDE2A,ENPP3
15_Summary	GO Biological Processes	GO:0043484	regulation of RNA splicing	-2.81	-0.39	4/141	1195,1196,6625,7072	CLK1,CLK2,SNRNP70,TIA1
15_Member	GO Biological Processes	GO:0043484	regulation of RNA splicing	-2.81	-0.39	4/141	1195,1196,6625,7072	CLK1,CLK2,SNRNP70,TIA1
16_Summary	GO Biological Processes	GO:0021872	forebrain generation of neurons	-2.73	-0.35	3/69	2260,55558,57688,1291,1742,2335,3690,3996,57580,80024,57007	FGFR1,PLXNA3,ZSWIM6,COL6A1,DLG4,FN1,ITGB3,LLGL1,PREX1,SLC8B1,ACKR3
16_Member	GO Biological Processes	GO:0021872	forebrain generation of neurons	-2.73	-0.35	3/69	2260,55558,57688	FGFR1,PLXNA3,ZSWIM6
16_Member	GO Biological Processes	GO:0000904	cell morphogenesis involved in differentiation	-2.29	-0.01	8/756	1291,1742,2335,3690,3996,55558,57580,57688	COL6A1,DLG4,FN1,ITGB3,LLGL1,PLXNA3,PREX1,ZSWIM6
16_Member	GO Biological Processes	GO:0050920	regulation of chemotaxis	-2.12	0.00	4/220	2260,55558,57688,80024	FGFR1,PLXNA3,ZSWIM6,SLC8B1
16_Member	GO Biological Processes	GO:0006935	chemotaxis	-2.10	0.00	7/649	2260,3690,55558,57007,57580,57688,80024	FGFR1,ITGB3,PLXNA3,ACKR3,PREX1,ZSWIM6,SLC8B1
16_Member	GO Biological Processes	GO:0042330	taxis	-2.09	0.00	7/651	2260,3690,55558,57007,57580,57688,80024	FGFR1,ITGB3,PLXNA3,ACKR3,PREX1,ZSWIM6,SLC8B1
16_Member	GO Biological Processes	GO:0010975	regulation of neuron projection development	-2.03	0.00	6/512	1742,2260,2335,55558,57580,57688	DLG4,FGFR1,FN1,PLXNA3,PREX1,ZSWIM6
17_Summary	GO Biological Processes	GO:0010876	lipid localization	-2.51	-0.17	6/405	1109,3690,57104,57198,57476,151176	AKR1C4,ITGB3,PNPLA2,ATP8B2,GRAMD1B,ERFE
17_Member	GO Biological Processes	GO:0010876	lipid localization	-2.51	-0.17	6/405	1109,3690,57104,57198,57476,151176	AKR1C4,ITGB3,PNPLA2,ATP8B2,GRAMD1B,ERFE
17_Member	GO Biological Processes	GO:0006869	lipid transport	-2.03	0.00	5/365	1109,3690,57198,57476,151176	AKR1C4,ITGB3,ATP8B2,GRAMD1B,ERFE
18_Summary	GO Biological Processes	GO:0060337	type I interferon signaling pathway	-2.34	-0.04	3/95	3105,3394,81030,2335,3500,3718,23529,57007	HLA-A,IRF8,ZBP1,FN1,IGHG1,JAK3,CLCF1,ACKR3

GroupID	Category	Term	Description	Log P	Log(q-value)	InTerm_InList	Genes	Symbols
18_Member	GO Biological Processes	GO:0060337	type I interferon signaling pathway	-2.34	-0.04	3/95	3105,3394,81030	HLA-A,IRF8,ZBP1
18_Member	GO Biological Processes	GO:0071357	cellular response to type I interferon	-2.34	-0.04	3/95	3105,3394,81030	HLA-A,IRF8,ZBP1
18_Member	GO Biological Processes	GO:0034340	response to type I interferon	-2.29	-0.01	3/99	3105,3394,81030	HLA-A,IRF8,ZBP1
18_Member	GO Biological Processes	GO:0019221	cytokine-mediated signaling pathway	-2.16	0.00	8/796	2335,3105,3394,3500,3718,23529,57007,81030	FN1,HLA-A,IRF8,IGHG1,JAK3,CLCF1,ACKR3,ZBP1
19_Summary	GO Biological Processes	GO:1903530	regulation of secretion by cell	-2.22	0.00	8/777	43,2260,2335,3996,10874,64780,80024,91683,5138,22994,26353,27111	ACHE,FGFR1,FN1,LLGL1,NMU,MICAL1,SLC8B1,SYT12,PDE2A,CEP131,HSPB8,SDCBP2
19_Member	GO Biological Processes	GO:1903530	regulation of secretion by cell	-2.22	0.00	8/777	43,2260,2335,3996,10874,64780,80024,91683	ACHE,FGFR1,FN1,LLGL1,NMU,MICAL1,SLC8B1,SYT12
19_Member	GO Molecular Functions	GO:0042803	protein homodimerization activity	-2.11	0.00	7/646	43,2260,5138,22994,26353,27111,80024	ACHE,FGFR1,PDE2A,CEP131,HSPB8,SDCBP2,SLC8B1
20_Summary	Reactome Gene Sets	R-HSA-8957275	Post-translational protein phosphorylation	-2.18	0.00	3/108	2335,57104,57719	FN1,PNPLA2,ANO8
20_Member	Reactome Gene Sets	R-HSA-8957275	Post-translational protein phosphorylation	-2.18	0.00	3/108	2335,57104,57719	FN1,PNPLA2,ANO8
20_Member	Reactome Gene Sets	R-HSA-381426	Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Bi	-2.01	0.00	3/125	2335,57104,57719	FN1,PNPLA2,ANO8

Supplemental Table 2 Metascape Enrichment analysis- Genes higher in subfertile mares

GroupID	Category	Term	Description	LogP	Log(q-value)	InTerm_InList	Genes	Symbols
1_Summary	GO Molecular Functions	GO:0008028	monocarboxylic acid transmembrane transporter activity	-5.57	-0.95	3/57	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0008028	monocarboxylic acid transmembrane transporter activity	-5.57	-0.95	3/57	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0015293	symporter activity	-4.34	-0.29	3/146	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0005342	organic acid transmembrane transporter activity	-4.30	-0.29	3/150	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0046943	carboxylic acid transmembrane transporter activity	-4.30	-0.29	3/150	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Biological Processes	GO:0015718	monocarboxylic acid transport	-4.15	-0.23	3/169	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0008514	organic anion transmembrane transporter activity	-3.85	-0.01	3/213	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0015291	secondary active transmembrane transporter activity	-3.71	0.00	3/239	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0008509	anion transmembrane transporter activity	-3.32	0.00	3/324	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Biological Processes	GO:0015849	organic acid transport	-3.28	0.00	3/334	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Biological Processes	GO:0046942	carboxylic acid transport	-3.28	0.00	3/334	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Molecular Functions	GO:0022804	active transmembrane transporter activity	-3.17	0.00	3/365	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Biological Processes	GO:0015711	organic anion transport	-2.81	0.00	3/487	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
1_Member	GO Biological Processes	GO:0006820	anion transport	-2.47	0.00	3/637	6555,9121,220963	SLC10A2,SLC16A5,SLC16A9
2_Summary	GO Biological Processes	GO:0006605	protein targeting	-2.97	0.00	3/426	6144,6218,54543	RPL21,RPS17,TOMM7
2_Member	GO Biological Processes	GO:0006605	protein targeting	-2.97	0.00	3/426	6144,6218,54543	RPL21,RPS17,TOMM7
2_Member	GO Biological Processes	GO:0072594	establishment of protein localization to organelle	-2.66	0.00	3/548	6144,6218,54543	RPL21,RPS17,TOMM7